


```

1 TELEFAX: (712) 474-6448
2
3 FILE# 25-8856
4
5 INFORMATION FOR SEQ ID NO: 40:
6
7 SEQUENCE CHARACTERISTICS:
8
9 LENGTH: 880 amino acids
10
11 TYPE: amino acid
12
13 TOPOLOGY: linear
14
15 M-TERM: TYPE: protein
16
17 SEQUENCE DESCRIPTION: SEQ ID NO: 40
18
19 US-09-088647H-40

```

Query Match	8.28;	Score	130.5;	DB	4;	Length	80;
Host local similarity	19.08;	Pred. No.	2.1e-05;				
Matches	66;	Conservative	79;	Mismatches	125;	In bits	77;
						Caps	16;

[illegible]

RESULT 4
 OCT-09-1154-40 Application P/TOS9-11153
 Sequence 40: Application P/TOS9-11153
 GENERAL INFORMATION:
 APPLICANT: Louis A. Bond
 APPLICANT: Gray, Patrick W.
 APPLICANT: Schwabart, Vicki L.
 TITLE OF INVENTION: Novel Seven Transmembrane Receptors
 NUMBER OF SEQUENCES: 64
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ADDRESSEE: Hickoll
 STREET: 6600 Sears Tower, 244 South Wacker Drive
 CITY: Chicago,
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: attached Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/11154
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/977,452
 FILING DATE: 17-NOV-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Noland, Greta E.
 REGISTRATION NUMBER: 45,402

1 REFERENCE: J. M. FELT, NIMR-3, 31794
 2 TELECOMMUNICATIONS INFORMATION:
 3 TELEPHONE: (312) 474-6400
 4 TELEFAX: (312) 474 0448
 5 FAX: 25-4856
 6 INFORMATION FOR SEQ ID NO: 40:
 7 SEQUENCE CHARACTERISTICS:
 8 LENGTH: 380 amino acids
 9 TYPE: amino acid
 10 COMPOSITION: linear
 11 MOLECULE TYPE: protein
 12 ECT-US94-11153-40

Query Match	8.2%	Score	130.5	DB %	Length	800			
Best Local Similarity	19.0%	Pred. No.	2.1e-05						
Matches	66	Conservative	79	Mismatches	125	Indels	77	Gaps	16

[illegible]

```

1  RESULT 5
2  US-08-570-157-7
3  Sequence 7, Application US/08570157
4  Patent No. 5750453
5  GENERAL INFORMATION:
6  APPLICANT: Koplin, Alan S.
7  APPLICANT: Reinhold, Martin
8  TITLE OF INVENTION: ASSAY FOR NEW PEPTIDE ANALOGS
9  TITLE OF INVENTION: PEPTIDE HORMONE RECEPTORS
10 NUMBER OF SEQUENCES: 23
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: Fish & Richardson P.C.
13 STREET: 225 Franklin Street
14 CITY: Boston
15 STATE: MA
16 COUNTRY: USA
17 ZIP: 02110-2804
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTING: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: Project Release #1.0, Version #1.00
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/08570157
25 FILING DATE: 11-06-1995
26 CLASSIFICATION: 435
27 ATE CRYSTALLINE INTERMEDIATE:
28 NAME: Clark, Paul L.
29 REGISTRATION NUMBER: 40,142
30 REFERENCE/OWNERS: MORGAN, DOUGLAS/109001

```

[illegible][illegible]


```

07 4 PLEPTESSLVVTVVTFVNFANFATVAVNSIFWEPKPT SPANQHTALAVS----- 64
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 41 PLEPTESSLVVTVVTFVNFANFATVAVNSIFWEPKPT SPANQHTALAVS----- 62
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
07 55 -----RVDLL-----WTLINWSTVAVNANFNSVEVHLAVNANWATNINININ 97
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 93 FFWYVYAAWVWVETLTKMWSMGLWVYSLFVLMMLD-----PVALVAVVAVSL 146
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
07 98 LATTLLVTVVTFVNFANFATVAVNSIFWEPKPT SPANQHTALAVS----- 157
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 147 RAPTIVVTVVTFVNFANFATVAVNSIFWEPKPT SPANQHTALAVS----- 191
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
07 158 FFWYVYAAWVWVETLTKMWSMGLWVYSLFVLMMLD-----PVALVAVVAVSL 215
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 192 SINSTWV-----LSSHTINTDGVTFVLMMLD-----PVALVAVVAVSL 248
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
07 216 SQPSTVYHAKALGVISPLD-----ALVETLSIMSVSPSLNKKVPMPT--- 264
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 244 -----KAVVMTLAVVAVVLEFWPVPNVLPLETEL-----EVLCVETFE 279
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
07 265 -----KATGSPSDFEFLWVNFQVETPSV 294
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 280 RYLVAVATVETLAVVMTNIPVTFVDFKPKVYTLQF 419
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

```

RESULT 8

```

US-09-045-644-56
Sequence 55, Application US/90A4564
Patent No. 6287805
GENERAL INFORMATION:
APPLICANT: Graham, Gerard J. et al.
TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & CORKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patoutin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/99/045, 563
FILING DATE: 20 MAR-98
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MANDADOURAS, Amy E.
REGISTRATION NUMBER: 46,207
REFERENCE/CHECK NUMBER: PNT-044
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
US-09-045-584-55

```

Query Match 7.78; Score 124; DB 4; Length 360;
 Best Local Similarity 18.58; Prod. No. 0.00012;
 Match 4.3; Conservative 6.2; Mismatches 105; Indels 110; Gaps 11.

```

07 4 PLEPTESSLVVTVVTFVNFANFATVAVNSIFWEPKPT SPANQHTALAVS----- 64
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 41 PLEPTESSLVVTVVTFVNFANFATVAVNSIFWEPKPT SPANQHTALAVS----- 62
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
07 55 -----RVDLL-----WTLINWSTVAVNANFNSVEVHLAVNANWATNINININ 97
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 93 FFWYVYAAWVWVETLTKMWSMGLWVYSLFVLMMLD-----PVALVAVVAVSL 146
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
07 98 LATTLLVTVVTFVNFANFATVAVNSIFWEPKPT SPANQHTALAVS----- 157
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 147 RAPTIVVTVVTFVNFANFATVAVNSIFWEPKPT SPANQHTALAVS----- 191
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
07 158 FFWYVYAAWVWVETLTKMWSMGLWVYSLFVLMMLD-----PVALVAVVAVSL 215
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 192 SINSTWV-----LSSHTINTDGVTFVLMMLD-----PVALVAVVAVSL 248
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
07 216 SQPSTVYHAKALGVISPLD-----ALVETLSIMSVSPSLNKKVPMPT--- 264
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 244 -----KAVVMTLAVVAVVLEFWPVPNVLPLETEL-----EVLCVETFE 279
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
07 265 -----KATGSPSDFEFLWVNFQVETPSV 294
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
10 280 RYLVAVATVETLAVVMTNIPVTFVDFKPKVYTLQF 419
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

```

RESULT 9

```

US-08-118-270-51
Sequence 51, Application US/98118270
Patent No. 5508484
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEWMARK
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: 5 1/4" disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patoutin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/98/118,270
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,044
REFERENCE/CHECK NUMBER: MURPHY 2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5192
TELEFAX: 202-747-4528
TELEX: 248634
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-51

```

Query Match 7.44; Score 117.5; DB 1; Length 325;
 Best Local Similarity 24.18; Prod. No. 0.00047;

[illegible]

[illegible]

Search completed: August 12, 2002, 13:50:04
 Job time: 2285 sec

Query Match: 1.99% Score 18; DB 4; Length 1982;
 Host Local Similarity: 100.0%; Prod. No. 14;
 Matches: 18; Conservation: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 DB: 1723 TATGAAATTTCCTAA 1706

RESULT: 4
 US-09-005-051-15
 Sequence 15, Application US/09005051
 Patent No. 6,291,222
 GENERAL INFORMATION:
 APPLICANT: SILVER, Gary W.
 APPLICANT: WISNOSKI, Nancy
 TITLE OF INVENTION: Nucleic acid
 TITLE OF INVENTION: Methods, products and uses thereof
 NUMBER OF SEQUENCES: 66
 ADDRESS: Carol Talkington Verser, Ph.D.
 ADDRESS: Carol Talkington Verser, Ph.D.
 STREET: 1825 Sharp Point Drive
 CITY: Fort Collins
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80525
 TELEPHONE: 970/493-7272
 TELEFAX: 970/493-7272
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1982 nucleotides
 TYPE: nucleic acid
 STRANDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 41..1517
 FEATURE:
 NAME/KEY: Amp
 LOCATION: 400
 US-09-005-051-13

Query Match: 1.99% Score 18; DB 4; Length 1982;
 Host Local Similarity: 100.0%; Prod. No. 14;
 Matches: 18; Conservation: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 DB: 1723 TATGAAATTTCCTAA 277
 RESULT: 4
 US-09-005-051-15
 Sequence 15, Application US/09005051
 Patent No. 6,291,222
 GENERAL INFORMATION:
 APPLICANT: SILVER, Gary W.
 APPLICANT: WISNOSKI, Nancy
 TITLE OF INVENTION: Nucleic acid
 TITLE OF INVENTION: Methods, products and uses thereof
 NUMBER OF SEQUENCES: 66
 ADDRESS: Carol Talkington Verser, Ph.D.
 ADDRESS: Carol Talkington Verser, Ph.D.
 STREET: 1825 Sharp Point Drive
 CITY: Fort Collins
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80525
 TELEPHONE: 970/493-7272
 TELEFAX: 970/493-7272
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1982 nucleotides
 TYPE: nucleic acid
 STRANDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 41..1517
 FEATURE:
 NAME/KEY: Amp
 LOCATION: 400
 US-09-005-051-13

Query Match: 1.99% Score 18; DB 4; Length 1982;
 Host Local Similarity: 100.0%; Prod. No. 14;
 Matches: 18; Conservation: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 DB: 1723 TATGAAATTTCCTAA 1706

RESULT: 4
 US-09-005-051-15
 Sequence 15, Application US/09005051
 Patent No. 6,291,222
 GENERAL INFORMATION:
 APPLICANT: SILVER, Gary W.
 APPLICANT: WISNOSKI, Nancy
 TITLE OF INVENTION: Nucleic acid
 TITLE OF INVENTION: Methods, products and uses thereof
 NUMBER OF SEQUENCES: 66
 ADDRESS: Carol Talkington Verser, Ph.D.
 ADDRESS: Carol Talkington Verser, Ph.D.
 STREET: 1825 Sharp Point Drive
 CITY: Fort Collins
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80525
 TELEPHONE: 970/493-7272
 TELEFAX: 970/493-7272
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1982 nucleotides
 TYPE: nucleic acid
 STRANDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 41..1517
 FEATURE:
 NAME/KEY: Amp
 LOCATION: 400
 US-09-005-051-13


```

: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 1410 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   MOLECULE TYPE: DNA (genomic)
:   HYPOTHETICAL: N
:   ANTI-SENSE: N
:   ORIGIN: +-----

```

```

Query Match      100%   Score 172   DB 3   Length 1410
Post local Similarity 100.0%   Prod. No. 41:
Matches 172   Conservation 0%   Mismatches 0%   Indels 0%   Gaps 0%

QY      4  GATACTTTCTGACCA 19
      1111111111111111
DB      623  GATAACTTTCTTACCA 639

```

Search completed: August 12, 2002, 18:13:07
Job time: 3949 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

CM nucleotide - nucleotide search, using SW model

Run on: August 12, 2002, 11:34:43 : Search time 49.18 Seconds

(Without alignments)
4644.961 Million cell updates/sec

Title: US-09-825-882-7

Percent score: 1.0
Sequence: 1.0
Database: 1.0

Scoring table: IDENTITY_NUC
Gapop: 10.0 : Gapext: 1.0

Search: 48343 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767866

Minimum hit seq length: 0
Maximum hit seq length: 2000000000

Post processing: Minimum Match 100%

Listing first 45 summaries

Database:

Issued_Patents_NA:
1. 1992-11-14-91 for FA_CMR seq 1
2. 1992-11-14-91 for FA_CMR seq 1
3. 1992-11-14-91 for FA_CMR seq 1
4. 1992-11-14-91 for FA_CMR seq 1
5. 1992-11-14-91 for FA_CMR seq 1
6. 1992-11-14-91 for FA_CMR seq 1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	40.4	4.3	5622	4	US-09-825-882-7
2	40.4	4.3	5622	4	US-09-825-882-7
3	36	3.9	7218	1	US-08-242-463-14
4	35.8	3.8	1510	1	US-08-248-466B-3
5	35.8	3.8	1512	1	US-08-248-466B-11
6	34.8	3.7	3288	1	US-08-208-008C-3
7	34	3.7	2989	3	US-08-433-522A-9
8	34	3.7	2989	3	US-09-115-166-9
9	34	3.7	2989	4	US-08-942-046-9
10	34	3.6	2629	2	US-09-012-030-1
11	34.4	3.5	2629	2	US-08-590-454-1
12	33	3.5	14602	1	US-08-597-236-1
13	33	3.5	14602	1	US-08-746-682A-1
14	32.4	3.5	2974	3	US-08-433-522A-7
15	32.4	3.5	2974	3	US-09-115-166-7
16	32.4	3.5	2974	4	US-08-942-046-7
17	32.4	3.4	940	1	US-08-448-600-2
18	32	3.4	3088	3	US-08-947-610-3
19	31.6	3.4	276	1	US-07-800-370-1
20	31.6	3.4	276	1	US-08-057-169-1
21	31.6	3.4	484	1	US-09-499-47
22	31.6	3.4	468	4	US-09-299-268-47
23	31.6	3.4	895	3	US-08-924-747-7
24	31.6	3.4	895	4	US-09-247-373B-7
25	31.6	3.4	895	4	US-09-296-715-7
26	31.6	3.4	4156	1	US-08-465-687A-1
27	31.6	3.4	4156	3	US-09-030-970-1

28	31.6	3.4	4156	5	PCT-US94-1184-1
29	31.6	3.4	14176	1	US-08-407-499-1
30	31.6	3.4	14176	1	US-08-407-499-14
31	31.6	3.4	14176	4	US-09-299-268-1
32	31.6	3.4	14176	4	US-09-299-268-14
33	31.4	3.4	840	1	US-08-688-609-1
34	31.4	3.4	840	3	US-09-602-842-1
35	31.4	3.4	1875	1	US-08-286-425A-1
36	31.4	3.4	3234	1	US-08-286-425A-7
37	30.8	3.3	458	1	US-09-141-000-4
38	30.8	3.3	2454	1	US-08-459-696-3
39	30.8	3.3	47948	4	US-09-251-645-11
40	30.6	3.3	2965	2	US-08-460-570-1
41	30.6	3.3	2965	2	US-08-460-570-2
42	30.6	3.3	2965	3	US-08-286-870A-1
43	30.6	3.3	2965	3	US-08-286-870A-2
44	30.4	3.3	791	2	US-07-645-196A-4
45	30.4	3.3	1568	4	US-09-286-691-1

ALIGNMENTS

RESULT 1

US-09-067-800-4

Sequence 3, Application US/09067800

Patent No. 6156024

GENERAL INFORMATION:

APPLICANT: Yandovsky, Martin F.

INVENTOR: Yandovsky, Martin F.

FILE NO: 1994-01-01

CLASS OF INVENTION: Seed dispersal

TYPE OF SEQUENCE: 24

SEQUENCE: Campbell & Flores LLP

ADDRESS: 400 La Jolla Village Drive Suite 700

CITY: San Diego

STATE: California

COUNTRY: United States

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM pc compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/067,800

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 41,815

REFERENCE/SEQUENCE NUMBER: P-02 2948

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 545-9901

TELEFAX: (619) 545-8949

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5622 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

FEATURE:

NAME/KEY: music feature

LOCATION: 1-5622

OTHER INFORMATION: /label

OTHER: 09-067-800-4

Query Match: 4.3% Score 40.4: DB 4: Length 5622
Post local similarity: 53.1% Pred. No. 0.059
Matches: 867, conservative 702, Indels 0, gaps 0

101 AATTTTATTCGCATTAAGCAGCCGATTCGCAGAGCTTAATTTCTTAAGAAAAAATCA 132

```
NAME/KBY: CDS
LOCATCN: 419., 147
US-08-248-466B-11
```

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

```

1  HS: 08-433 522A-9
2  SOURCE: 9, Applicant bio 05/08415722A
3  Patent No. 403514
4  GENERAL INFORMATION:
5  APPLICANT: 36 INC, Palo
6  APPLICANT: THOMAS, Wayne
7  APPLICANT: YANO, Yui Hiro
8  APPLICANT: DOEMERER, Steven
9  APPLICANT: SIA, Two Yuan Charles
10 APPLICANT: KEVIN, Michael
11 TITLE OF INVENTION: HAEMOPHILUS GUTTER MEMBRANE PROTEIN
12 NUMBER OF SEQUENCES: 45
13 CORRESPONDENT ADDRESS:
14 ADDRESSEE: 414 N. M. BARRY
15 STREET: 614 Flood, 430 University Avenue
16 CITY: Toronto
17 STATE: Ontario
18 COUNTRY: Canada
19 ZIP: M5G 1B7
20
21 COMPUTER READABLE FORM:
22 MEDIUM TYPE: floppy disk
23 OPERATING SYSTEM: IBM PC compatible
24 SOFTWARE: Patent in Release #1.0, Version #1.25
25 CURRENT APPLICATION DATA:
26 APPLICATION NUMBER: 3758473, 522A
27 FILING DATE: 12 SEP 1995
28 CLASSIFICATION: 445
29
30 ATTORNEY/AGENT INFORMATION:
31 NAME: STEWART, Michael I
32 REGISTRATION NUMBER: 24,974
33 REFERENCE/KEY NUMBER: 1048 434 M15:10
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: (416) 595-1155
36 TELEFAX: (416) 595-1164
37 INFORMATION FOR SEQ ID NO: 9:
38 SEQUENCE CHARACTERISTICS:
39 LENGTH: 2969 base pairs
40 TYPE: nucleic acid
41 STRANDEDNESS: single
42 TOPOLOGY: linear
43 MOLECULE TYPE: DNA (genomic)
44 FEATURE:
45 NAME/KEY: CDS
46 LOCATION: 433..2708
47
48 08-433 522A-9

```

[illegible]

RESULT 8
 US-09-145-166-9
 ; Sequence 9, Application US/09145166
 ; Patent No. 6083744

Module	Type	DNA (genomic)
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	452..1803	
OTHER INFORMATION:	/product "cpsA"	
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	1907..2545	
OTHER INFORMATION:	/product "cpsB"	
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	2547..3239	
OTHER INFORMATION:	/product "cpsC"	
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	3244..3995	
OTHER INFORMATION:	/product "cpsD"	
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	4051..4741	
OTHER INFORMATION:	/product "cpsE"	
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	4898..5854	
OTHER INFORMATION:	/product "cpsF"	
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	6425..7540	
OTHER INFORMATION:	/product "cpsG"	
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	7746..8212	
OTHER INFORMATION:	/product "cpsH"	
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	8221..9192	
OTHER INFORMATION:	/product "cpsI"	
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	9285..10364	
OTHER INFORMATION:	/product "cpsJ"	
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	10392..11339	
OTHER INFORMATION:	/product "cpsK"	
FEATURES:		
NAME/KEY:	misc_feature	
LOCATION:	11402..12232	
OTHER INFORMATION:	/product "CDS (cps L) coding CDS"	
OTHER INFORMATION:	(cps K) on nucleotides 10392-11339"	
FEATURES:		
NAME/KEY:	CDS	
LOCATION:	12243..13651	
OTHER INFORMATION:	/product "cpsM"	
FEATURES:		
NAME/KEY:	misc_feature	
LOCATION:	13742..14305	
OTHER INFORMATION:	/product "CDS on the"	
OTHER INFORMATION:	complementary strand"	
OTHER INFORMATION:	/product "orf7"	
FEATURES:		
NAME/KEY:	terminator	
LOCATION:	246..252	
FEATURES:		
NAME/KEY:	promoter	
LOCATION:	274..302	
FEATURES:		
NAME/KEY:	PRS	
LOCATION:	440..445	

7. I did not know that the Government was not going to be able to pay the interest on the loan.

therapy. Pathological conditions can be diagnosed by determining the

100

DK N-PSDB: AAF92515.
 XX
 XX Nucleic acids encoding the 12R family of G-protein coupled taste
 PT receptors, useful for identifying taste modulators that can be used in
 PI food and pharmaceutical industries to customize taste, for e.g., to
 XX decrease the bitter taste of food -
 PS Claim 19: Page 172: 24pp: English.
 XX
 XX AAF92512 to AAF92572 represent nucleic acids which encode taste
 CC transduction G-protein coupled receptors designated 12R proteins.
 CC AAB87741 to AAB87824 represent 12R proteins, and AAB87825 to AAB87840
 CC represent 12R family consensus sequences from the present invention.
 CC The 12R proteins are taste modulators. The nucleic acids are useful as
 CC probes for the identification of taste cells, as the nucleic acids are
 CC specifically expressed in taste cells. They also serve as tools for the
 CC generation of taste topographic maps that elucidate the relationship
 CC between the taste cells of the tongue and taste sensory neurons, leading
 CC to taste centers in the brain. The taste modulators are useful for
 CC pharmacological and genetic modulation of taste signaling pathways.
 CC Modulatory compounds comprising 12R proteins can therefore be used to
 CC food and pharmaceutical industries to customize taste, for e.g., to
 CC decrease the bitter taste of food -
 XX
 XX Sequence: 299 AA:

Query Match 8.1% Score 25: DB 22: Length 299:
 Best Local Similarity 100.0%: Prod. No. 4,96-15:
 Matches 25: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 57 37 KKKKISFANDQIFAIASVSWLW 61
 10 37 KKKKISFANDQIFAIASVSWLW 61
 10 37 KKKKISFANDQIFAIASVSWLW 61

RESULT 12
 AAM23537
 ID AAM23537 standard: Protein: 148 AA.
 AC AAM23537:
 XX
 XX 12 OCT-2001 (first entry)
 XX
 XX Human EST encoded protein SEQ ID NO: 1062.
 CC
 CC Human sheep pit. cow. fruit fly. yeast. bacteria. mycoplasma. horse.
 CC tomato. monkey. dog. sea urchin. expressed sequence tag. EST.
 CC diagnostics: forensic test; gene mapping; genetic disorder;
 CC biodiversity; gene therapy; mutation.
 XX
 XX Homo sapiens:
 CC
 CC W0200154477-A2.
 XX
 XX 02-AUG-2001.
 XX
 XX 25-JAN-2001: 2001W0-U02687.
 XX
 XX 25-JAN-2001: 2003US-0491404.
 CC 17-SEP-2000: 2000US-0617746.
 CC 03-AUG-2000: 2000US-0641451.
 CC 15-SEP-2000: 2000US-0661870.
 XX
 XX (HENSE) HENSE INC.
 XX
 XX Tang YF, Liu Y, Zhou P, Qian XB, Wang Z, Chen R, Asanul V,
 PI Gao Y, Jinnah KA, Zhang J, Weinman J:
 XX WPI: 2001 476194/51
 CC N-PSDB: AAB98196.
 XX
 XX Isolated polypeptide for treatment of diseases, diagnostics, raising

PT antibodies and research use -
 XX
 XX Claim 20: Page 802: 127pp: English.
 CC
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, mycoplasma, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a
 CC protein of the invention.
 XX
 XX Sequence: 148 AA:

Query Match 7.8% Score 24: DB 22: Length 148:
 Best Local Similarity 100.0%: Prod. No. 1,86-14:
 Matches 24: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 57 107 IIPSTIYHFFPMIIPFQYQNS 220
 26 107 IIPSTIYHFFPMIIPFQYQNS 220

RESULT 13
 AAB87750
 ID AAB87750 standard: Protein: 129 AA.
 AC AAB87750:
 XX
 XX 16-MAY-2001 (first entry)
 XX
 XX Human: 12R20 amino acid sequence SEQ ID NO:37.
 CC
 CC Human: rat; mouse; 12R: taste receptor; G-protein coupled receptor;
 CC taste transduction G-protein coupled receptor; identification, tongue;
 CC taste sensory neuron; taste cell; taste modulator; food;
 CC taste signaling pathway.
 XX
 XX Homo sapiens:
 CC
 CC W0200118050-A2.
 XX
 XX 15-MAR-2001.
 XX
 XX 09-SEP-2000: 2000W0-U024821.
 CC 10-SEP-1999: 99US-0494644.
 CC 22-FEB-2000: 2000US-0510342.
 XX
 XX (BAC) ONIV CALIFORNIA,
 CC (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Zaker CS, Adler JJ, Ryba N, Mueller K, Boon M:
 XX WPI: 2001-211396/21.
 CC N-PSDB: AAF92519.
 XX
 XX Nucleic acids encoding the 12R family of G-protein coupled taste
 PT receptors, useful for identifying taste modulators that can be used in
 PI food and pharmaceutical industries to customize taste, for e.g., to
 XX decrease the bitter taste of food -
 PS Claim 19: Page 176: 24pp: English.
 XX
 XX AAF92502 to AAF92572 represent nucleic acids which encode taste
 CC transduction G-protein coupled receptors designated 12R proteins.
 CC AAB87741 to AAB87824 represent 12R proteins, and AAB87825 to AAB87840
 CC represent 12R family consensus sequences from the present invention.
 CC The 12R proteins are taste modulators. The nucleic acids are useful as
 CC probes for the identification of taste cells, as the nucleic acids are
 CC specifically expressed in taste cells. They also serve as tools for the





[illegible][illegible]

D1 16-MAY-2001 (first entry)
 XX Human T2R37 nucleotide sequence SPO ID NO:63
 FE Human rat: mouse: 12R: taste receptor: G-protein coupled receptor:
 XX taste sensory neuron: taste cell: taste modulator: food:
 KM taste sensory neuron: taste cell: taste modulator: food:
 KM taste signaling pathway: ds.
 XX Homo sapiens.
 OS
 FN W0200118050-A7.
 XX
 FE 15-MAR-2001.
 XX
 FE 08-SEP-2000: 2000W-0824821.
 XX
 FE 10-SEP-1999: 9908-0393634.
 XX
 FE 22-FEB-2000: 2000US-0510382.
 XX
 PA (RSC)) UNIV CALIFORNIA
 PA (USNH)) US DEPT HEALTH & HUMAN SERVICES.
 XX
 P1 Zaker SE, Adler JE, Pyda N, Mueller K, Heen M:
 P1 WPI: 2001-211396/21.
 P1 P-PSDB: AAB87767.
 XX
 FE Nucleic acids encoding the 12R family of G-protein coupled taste
 P1 receptors, useful for identifying taste modulators that can be used in
 P1 food and pharmaceutical industries to customize taste, for e.g. to
 P1 decrease the bitter taste of food -
 P1
 P1 Disclosure: Page 185: 249pp: English.
 XX
 FE AAF92502 to AAF92572 represent nucleic acids which encode taste
 P1 transduction G-protein coupled receptors designated 12R proteins
 P1 AAB87731 to AAB87824 represent 12R proteins, and AAB87825 to AAB87830
 P1 represent 12R family consensus sequences from the present invention.
 P1 The 12R proteins are taste modulators. The nucleic acids are useful as
 P1 probes for the identification of taste cells, as the nucleic acids are
 P1 specifically expressed in taste cells. They also serve as tools for the
 P1 generation of taste topographic maps that elucidate the relationship
 P1 between the taste cells of the tongue and taste sensory neurons leading
 P1 to taste centres in the brain. The taste modulators are useful for
 P1 pharmacological and genetic modulation of taste signalling pathways.
 P1 Modulatory compounds comprising 12R proteins can therefore be used in
 P1 food and pharmaceutical industries to customize taste, for e.g. to
 P1 decrease the bitter taste of food or drugs.
 XX
 SE Sequence 960 BP: 254 A: 178 C: 181 G: 417 T: 0 other:
 XX
 Query Match: 5.98; Score 55; DB 22; Length 960;
 Best local similarity 100.0%; Pred. No. 1,20-16;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 FE 640 aaagatctcaaatccaaatccaaatccatataaagctttcaaatatga 694
 FE ||||||||||||||||||||||||||||||||||||||||||||||||
 FE 640 aaagatctcaaatccaaatccaaatccatataaagctttcaaatatga 694
 XX
 RESULT 14
 AAF92524
 ID AAF92524 standard; DNA: 960 BP.
 XX
 FE AAF92524:
 XX
 FE 16-MAY-2001 (first entry)
 XX
 FE Human T2R30 nucleotide sequence SPO ID NO:52.
 XX
 KM Human rat: mouse: 12R: taste receptor: G-protein coupled receptor:

KM taste transduction G-protein coupled receptor: identification: tongue;
 KM taste sensory neuron: taste cell: taste modulator: food;
 KM taste signaling pathway: ds.
 XX
 OS Homo sapiens.
 FN W0200118050-A7.
 XX
 FE 15-MAR-2001.
 XX
 FE 08-SEP-2000: 2000W-0824821.
 XX
 FE 10-SEP-1999: 9908-0393634.
 XX
 FE 22-FEB-2000: 2000US-0510382.
 XX
 PA (RSC)) UNIV CALIFORNIA
 PA (USNH)) US DEPT HEALTH & HUMAN SERVICES.
 XX
 P1 Zaker SE, Adler JE, Pyda N, Mueller K, Heen M:
 P1 WPI: 2001-211396/21.
 P1 P-PSDB: AAB87760.
 XX
 FE Nucleic acids encoding the 12R family of G-protein coupled taste
 P1 receptors, useful for identifying taste modulators that can be used in
 P1 food and pharmaceutical industries to customize taste, for e.g. to
 P1 decrease the bitter taste of food -
 P1
 P1 Disclosure: Page 180-181: 249pp: English.
 XX
 FE AAF92562 to AAF92572 represent nucleic acids which encode taste
 P1 transduction G-protein coupled receptors designated 12R proteins.
 P1 AAB87731 to AAB87824 represent 12R proteins, and AAB87825 to AAB87830
 P1 represent 12R family consensus sequences from the present invention.
 P1 The 12R proteins are taste modulators. The nucleic acids are useful as
 P1 probes for the identification of taste cells, as the nucleic acids are
 P1 specifically expressed in taste cells. They also serve as tools for the
 P1 generation of taste topographic maps that elucidate the relationship
 P1 between the taste cells of the tongue and taste sensory neurons leading
 P1 to taste centres in the brain. The taste modulators are useful for
 P1 pharmacological and genetic modulation of taste signalling pathways.
 P1 Modulatory compounds comprising 12R proteins can therefore be used in
 P1 food and pharmaceutical industries to customize taste, for e.g. to
 P1 decrease the bitter taste of food or drugs.
 XX
 SE Sequence 960 BP: 256 A: 197 C: 183 G: 424 T: 0 other:
 XX
 Query Match: 5.98; Score 55; DB 22; Length 960;
 Best local similarity 100.0%; Pred. No. 1,20-16;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 FE 640 aaagatctcaaatccaaatccaaatccatataaagctttcaaatatga 694
 FE ||||||||||||||||||||||||||||||||||||||||||||||||
 FE 640 aaagatctcaaatccaaatccaaatccatataaagctttcaaatatga 694
 XX
 RESULT 15
 AAF92515
 ID AAF92515 standard; DNA: 960 BP.
 XX
 FE AAF92515:
 XX
 FE 16-MAY-2001 (first entry)
 XX
 FE Human T2R15 nucleotide sequence SPO ID NO:29.
 XX
 KM Human rat: mouse: 12R: taste receptor: G-protein coupled receptor:
 KM taste transduction G-protein coupled receptor: identification: tongue;
 KM taste sensory neuron: taste cell: taste modulator: food;
 KM taste signaling pathway: ds.
 XX
 OS Homo sapiens.

Genome version 4.5
Copyright (c) 1993 - 2000 CompuLink

0M protein - protein search, using sw model

Run on: August 12, 2002, 13:13:14 ; Search time 41.83 seconds
(without alignments)
677,427 Million cell updates/sec

Title: US-09-825-882-8
Percent score: 1597
Sequence: 1 MTEPLPILPSSIVVVFYVIG.....FLSWFQGMQRYWRCSEKTSNP 309

Scoring table:
Gapop 10.0 ; Gapext 0.5

Searched: 283108 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283108

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match: 0%
Listing first 45 summaries

Database: PIR_71: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Prod. No. is the number of results produced by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY:

Result No.	Score	Query Match	Length DB	ID	Description
1	130.5	8.2	341	2	126920
2	130.5	8.2	380	2	138435
3	126.5	7.3	387	2	082889
4	123	7.7	362	2	A57160
5	120.5	7.6	441	2	A70346
6	117	7.3	460	2	252587
7	116.5	7.3	415	2	869659
8	115.5	7.2	464	2	352543
9	115	7.2	411	2	A57223
10	115	7.2	458	2	111165
11	114.5	7.2	425	2	127852
12	112.5	7.0	412	2	A46750
13	112	7.0	355	2	132241
14	112	7.0	1898	2	T39039
15	111.5	7.0	464	2	148261
16	111.5	7.0	505	2	088206
17	111	7.0	349	2	123638
18	110.5	6.9	349	2	T25630
19	110.5	6.9	363	2	A49092
20	110.5	6.9	366	2	111286
21	110.5	6.9	548	2	064164
22	109	6.8	432	2	075622
23	108.5	6.8	387	2	855550
24	108.5	6.8	406	2	855549
25	108.5	6.8	770	2	G72589
26	108	6.8	380	2	052484
27	107.5	6.7	317	2	T33861
28	107.5	6.7	439	2	T25180
29	106.5	6.7	268	2	T16851

30	106.5	6.7	352	2	A45747	neurotrophin Y/pep
31	106.5	6.7	359	2	T51772	angiotensin II rec
32	106.5	6.7	906	2	096601	immunoglobulin 200
33	106	6.6	328	2	045774	odorant receptor 8
34	106	6.6	380	2	A48227	kappa opioid recep
35	106	6.6	380	2	S36143	kappa opioid recep
36	106	6.6	707	2	086894	hypothetical prote
37	105.5	6.6	274	2	088690	protein 141108.1
38	105.5	6.6	359	2	051104	angiotensin II rec
39	105	6.6	310	2	T22013	hypothetical prote
40	105	6.6	380	2	052488	kappa opioid recep
41	105	6.6	497	2	S59103	NAAD dehydrogenase
42	104.5	6.5	355	2	105027	G protein coupled
43	104.5	6.5	488	2	T29464	G protein coupled
44	104.5	6.5	477	2	125846	hypothetical prote
45	104	6.5	352	2	G00048	hypothetical prote
						tusin (LSTRA)

ALIGNMENTS

RESULT 1	126920	hypothetical protein Y45P108.6 - Caenorhabditis elegans	
C1Species:	Caenorhabditis elegans		
C2Date:	15 Oct 1999	Sequence revision 15 Oct 1999	Local change 21 Jul 2000
C3Accession:	126920		
Reference:	A.	Submitted to the EMBL Data Library, January 1998	
A1Reference number:	220286		
A2Accession:	126920		
A3Status:	Preliminary	Translated from GB/EMBL/DBD	
A4Molecule type:	DNA		
A5Residues:	1-441	<WIL>	
A6Cross reference:	EMBL:AL021487; JIDN:CAAL6458.1; GSPB8:0800022; CESP:Y45P108.6		
A7Experimental source:	clone Y45P108		
C1Genetics:			
A8Gene:	CESP:Y45P108.6		
A9Map position:	4		
A10Date:	1999		
C2Superfamily:	Caenorhabditis elegans hypothetical protein C4AL2.9b		
Query Match	Best local similarity	8.2%	Score 130.5; ID: 2; Length 341;
Matches	71; Conservative	54;	Mismatches 86; Indels 111; Gaps 20;
Q1	1	MTEPLPILPSSIVVVFYVIG	ALVSLHW
Q2	2	138435	ALVSLHW
Q3	3	082889	ALVSLHW
Q4	4	A57160	ALVSLHW
Q5	5	A70346	ALVSLHW
Q6	6	252587	ALVSLHW
Q7	7	869659	ALVSLHW
Q8	8	352543	ALVSLHW
Q9	9	A57223	ALVSLHW
Q10	10	111165	ALVSLHW
Q11	11	127852	ALVSLHW
Q12	12	A46750	ALVSLHW
Q13	13	132241	ALVSLHW
Q14	14	T39039	ALVSLHW
Q15	15	148261	ALVSLHW
Q16	16	088206	ALVSLHW
Q17	17	123638	ALVSLHW
Q18	18	T25630	ALVSLHW
Q19	19	A49092	ALVSLHW
Q20	20	111286	ALVSLHW
Q21	21	064164	ALVSLHW
Q22	22	075622	ALVSLHW
Q23	23	855550	ALVSLHW
Q24	24	855549	ALVSLHW
Q25	25	G72589	ALVSLHW
Q26	26	052484	ALVSLHW
Q27	27	T33861	ALVSLHW
Q28	28	T25180	ALVSLHW
Q29	29	T16851	ALVSLHW

1. *Chlorophyll a* (Chl *a*)

[illegible]

1. 2000年12月31日，本公司应收账款账面余额为1,000,000.00元，坏账准备余额为100,000.00元。

DB 1061 ENINSTRKIVRKGQTHRSQYTHRSFSTKQMSHSHHLLVYAWAWYRQSTNQT 1120
 A:Reference number: 221815
 A:Accession: 139009
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 11868 (GEN)
 A:Cross-references: EMBL/208543; FIRM/041064.1; SDBS/000066; SDBS/SP0612.076
 A:Experimental source: strain 9726; cosmid 16812
 A:Gene: SP0612.076
 A:Map position: 1
 A:Superfamily: Schizosaccharomyces hypodermal protein SP0612.076

RESULT 14

Query Match: 7.0%, Score 112, Db 2, Length 455
 Best local similarity: 24.7%, Prod. No. 0.44
 Matches: 74; Conservative: 46; Mismatches: 86; Gaps: 90; Gaps: 16
 DB 1061 ENINSTRKIVRKGQTHRSQYTHRSFSTKQMSHSHHLLVYAWAWYRQSTNQT 1120
 A:Reference number: 221815
 A:Accession: 139009
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 11868 (GEN)
 A:Cross-references: EMBL/208543; FIRM/041064.1; SDBS/000066; SDBS/SP0612.076
 A:Experimental source: strain 9726; cosmid 16812
 A:Gene: SP0612.076
 A:Map position: 1
 A:Superfamily: Schizosaccharomyces hypodermal protein SP0612.076

DB 1061 ENINSTRKIVRKGQTHRSQYTHRSFSTKQMSHSHHLLVYAWAWYRQSTNQT 1120
 A:Reference number: 221815
 A:Accession: 139009
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 11868 (GEN)
 A:Cross-references: EMBL/208543; FIRM/041064.1; SDBS/000066; SDBS/SP0612.076
 A:Experimental source: strain 9726; cosmid 16812
 A:Gene: SP0612.076
 A:Map position: 1
 A:Superfamily: Schizosaccharomyces hypodermal protein SP0612.076

RESULT 15

Query Match: 7.0%, Score 112, Db 2, Length 455
 Best local similarity: 24.7%, Prod. No. 0.44
 Matches: 74; Conservative: 46; Mismatches: 86; Gaps: 90; Gaps: 16
 DB 1061 ENINSTRKIVRKGQTHRSQYTHRSFSTKQMSHSHHLLVYAWAWYRQSTNQT 1120
 A:Reference number: 221815
 A:Accession: 139009
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 11868 (GEN)
 A:Cross-references: EMBL/208543; FIRM/041064.1; SDBS/000066; SDBS/SP0612.076
 A:Experimental source: strain 9726; cosmid 16812
 A:Gene: SP0612.076
 A:Map position: 1
 A:Superfamily: Schizosaccharomyces hypodermal protein SP0612.076

Query Match: 7.0%, Score 112, Db 2, Length 455
 Best local similarity: 24.7%, Prod. No. 0.44
 Matches: 74; Conservative: 46; Mismatches: 86; Gaps: 90; Gaps: 16
 DB 1061 ENINSTRKIVRKGQTHRSQYTHRSFSTKQMSHSHHLLVYAWAWYRQSTNQT 1120
 A:Reference number: 221815
 A:Accession: 139009
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 11868 (GEN)
 A:Cross-references: EMBL/208543; FIRM/041064.1; SDBS/000066; SDBS/SP0612.076
 A:Experimental source: strain 9726; cosmid 16812
 A:Gene: SP0612.076
 A:Map position: 1
 A:Superfamily: Schizosaccharomyces hypodermal protein SP0612.076

Query Match: 7.0%, Score 112, Db 2, Length 455
 Best local similarity: 24.7%, Prod. No. 0.44
 Matches: 74; Conservative: 46; Mismatches: 86; Gaps: 90; Gaps: 16
 DB 1061 ENINSTRKIVRKGQTHRSQYTHRSFSTKQMSHSHHLLVYAWAWYRQSTNQT 1120
 A:Reference number: 221815
 A:Accession: 139009
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 11868 (GEN)
 A:Cross-references: EMBL/208543; FIRM/041064.1; SDBS/000066; SDBS/SP0612.076
 A:Experimental source: strain 9726; cosmid 16812
 A:Gene: SP0612.076
 A:Map position: 1
 A:Superfamily: Schizosaccharomyces hypodermal protein SP0612.076

RESULT 15

Query Match: 7.0%, Score 112, Db 2, Length 455
 Best local similarity: 24.7%, Prod. No. 0.44
 Matches: 74; Conservative: 46; Mismatches: 86; Gaps: 90; Gaps: 16
 DB 1061 ENINSTRKIVRKGQTHRSQYTHRSFSTKQMSHSHHLLVYAWAWYRQSTNQT 1120
 A:Reference number: 221815
 A:Accession: 139009
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 11868 (GEN)
 A:Cross-references: EMBL/208543; FIRM/041064.1; SDBS/000066; SDBS/SP0612.076
 A:Experimental source: strain 9726; cosmid 16812
 A:Gene: SP0612.076
 A:Map position: 1
 A:Superfamily: Schizosaccharomyces hypodermal protein SP0612.076



Genome version 4.5
Copyright (c) 1993-2000 Genomics Ltd.

0M protein - protein search, using SW model

Run on: August 12, 2002, 13:57:53, Search time 19.71 seconds
(without alignments)

1506,423 Million cell updates/sec

Title: us-09-825-882-8

Percent score: 309
Sequence: 1 MTRFLPITFSSIVVVFVIV.....LISVWQKRWKRLISP 309

Sorting table:

cdlcp 60.0, cdepxt 60.0

Searched: 284138 seqs, 9608934 residues

Word size: 0

Total number of hits satisfying chosen parameters: 284138

Minimum (8 seq length) 0

Maximum (8 seq length) 200000000

Post processing: listing first 45 summaries

Database: 1: PIR-711*

2: PIR11*

3: PIR12*

4: PIR13*

5: PIR14*

6: PIR15*

7: PIR16*

8: PIR17*

9: PIR18*

10: PIR19*

11: PIR20*

12: PIR21*

13: PIR22*

14: PIR23*

15: PIR24*

16: PIR25*

17: PIR26*

18: PIR27*

19: PIR28*

20: PIR29*

21: PIR30*

22: PIR31*

23: PIR32*

24: PIR33*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	8	2.9	411	2	CH6638		glycosyl transferase
2	8	2.6	484	2	CH4955		hyp-N-acetylglutamate
3	7	2.3	83	2	C70073		hypothetical prote
4	7	2.3	106	2	H89982		truncated map w pr
5	7	2.3	120	2	S06731		lg kappa chain pro
6	7	2.3	120	2	S06732		lg kappa chain pro
7	7	2.3	128	2	S52448		lg kappa chain V I
8	7	2.3	128	2	J10073		aberrant kappa tra
9	7	2.3	131	1	KVMSM6		lg kappa chain pro
10	7	2.3	141	2	P01339		lg kappa chain pro
11	7	2.3	141	2	S55027		lg kappa chain pro
12	7	2.3	142	2	KVMS72		lg kappa chain pro
13	7	2.3	147	2	S55364		lg kappa chain pro
14	7	2.3	140	2	P80446		lg kappa chain pro
15	7	2.3	159	2	T23149		hypothetical prote
16	7	2.3	164	2	AH1727		transcription term
17	7	2.3	223	2	I46018		hypothetical prote
18	7	2.3	228	2	G86882		hypothetical prote
19	7	2.3	246	2	T13529		hypothetical prote
20	7	2.3	257	2	T25387		hypothetical prote
21	7	2.3	284	2	AH0726		phosphotransferase
22	7	2.3	284	2	S66427		antipeptidase (HC
23	7	2.3	286	1	W0RCMM		phosphotransferase
24	7	2.3	286	2	A98945		mannose specific p
25	7	2.3	286	2	D85793		PTS enzyme IID, ma
26	7	2.3	286	2	C81655		PTS enzyme IID, ma
27	7	2.3	295	2	AC0213		PTS system, mann
28	7	2.3	302	2	F36722		probable transmem
29	7	2.3	330	2	H70849		hypothetical prote

40	7	2.3	350	2	AG2476		permease protein p
41	7	2.3	371	2	H85536		CH68 hypothetical
42	7	2.3	371	2	E72097		conserved hypothet
43	7	2.3	376	2	T24223		hypothetical prote
44	7	2.3	393	2	G86330		hypothetical prote
45	7	2.3	396	2	I97183		hypothetical prote
36	7	2.3	402	2	J92164		hypothetical prote
37	7	2.3	402	2	I42615		hypothetical prote
38	7	2.3	410	2	F87004		probable glycopro
39	7	2.3	437	2	J03498		hypothetical prote
40	7	2.3	438	2	A98161		hypothetical prote
41	7	2.3	438	2	H86006		hypothetical prote
42	7	2.3	438	2	AH0997		high affinity ribo
43	7	2.3	439	2	S56491		ant-11 system to d
44	7	2.3	445	2	E22845		hypothetical prote
45	7	2.3	454	2	D86793		drug export prote

ALIGNMENTS

RESULT 1
CH6638
glycosyl transferase [imported], *Lactococcus lactis* subsp. *lactis* (strain 11.140)
C/Spectro: *Lactococcus lactis* subsp. *lactis*
C/Date: 22 Mar 2001 #Seqgeny_revision 23 Mar 2001 #FASTA_Change 03-Apr-2001
C/Accession: CH6638
Reference: A: Weicker, P.; Mauer, S.; Mallon, O.; Maloney, K.; Weissendorff, J.; Eh
Genome 808: 11, 741 753, 2001
A/Title: The complete genome sequence of the lactococcal bacterium *Lactococcus*
A/Reference number: AB6625; M010:2125186; M010:1134741
A/Accession: CH6638
A/Status: preliminary
A/Molecule type: DNA
A/Accession: 1-311-5870
A/Accession: CH-AP000179; F10:91272957; P108:AK03405; J1: C8990:CN00146
A/Experimental source: strain 11.1404
C/Genetics:
A/Genes: ybaI
C/Superfamily: stress response protein *csbB*

Query Match
Post local similarity 2.68; Score 8; DB 2; Length 411;
Matches 8; conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 102 LSTYLLK 109
DB 244 LSTYLLK 261
RESULT 2
CH4955
204 R acyl-hydrolytic activity [trans-3,3',2,2'-tetraol] *Escherichia coli* strain A
C/Spectro: *Escherichia coli* strain A
C/Date: 02 Mar 2001 #Seqgeny_revision 23 Mar 2001 #FASTA_Change 03-Apr-2001
C/Accession: CH4955
Reference: A: Maloney, K.; Mauer, S.; Mallon, O.; Maloney, K.; Weissendorff, J.; Eh
Nature 407: 81-86, 2000
A/Title: Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp.
A/Reference number: AB6430; M010:2045173
A/Accession: CH4955
A/Status: preliminary
A/Molecule type: DNA
A/Accession: 1-484-570
A/Cross references: CH-AP000179; C8990:CN00144
A/Experimental source: strain AFS
C/Genetics:
A/Genes: muref, M0215
A/Superfamily: hyp-N-acetylglutamate-aldolase ligase
C/Keywords: database

Query Match: 2.38; Score 7; DB 2; Length 140;
 Best Local Similarity: 100.00; Prod. No. 47;
 Matches: 7; conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 QY: 58 ELWILL: 64
 |||||
 10 7 ELWILL: 13

RESULT 15

12414
 hypothetical protein JCR.8 - Caenorhabditis elegans
 c-Spectra: Caenorhabditis elegans
 c-Date: 15-sep-1999 #sequence_revision 15-sep-1999 #loci_change 21-Jan-2000
 c-Accession: 12414
 c-Relidation: J.
 submitted to the EMBL data library: November 1996
 A-Reference number: Z19699
 A-Accession: 12414
 A-Status: preliminary; translated from cDNA/EMBL/DDNJ
 A-Molecule type: DNA
 A-Residues: 1359 SWISS
 A-Cross references: EMBL:Z92274; F11N:CAH5225.1; GSPB:GN0022; CESP:JCR.8
 A-Experimental source: clone JCR
 c-Cloned: CESP:JCR.8
 A-Name: CESP:JCR.8
 A-Map position: 4
 A-Introns: 8/27.4+/3
 c-Superfamily: Caenorhabditis hypothetical protein CA0H1.5

Query Match: 2.38; Score 7; DB 2; Length 159;
 Best Local Similarity: 100.00; Prod. No. 42;
 Matches: 7; conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 QY: 244 ELIATA: 240
 |||||
 10 26 ELIATA: 32

Search completed: August 12, 2002, 14:01:29
 Job time: 210 sec



10. 100% MATCH: 100% (100%)
 11. 100% MATCH: 100% (100%)
 12. 100% MATCH: 100% (100%)
 13. 100% MATCH: 100% (100%)
 14. 100% MATCH: 100% (100%)
 15. 100% MATCH: 100% (100%)
 16. 100% MATCH: 100% (100%)
 17. 100% MATCH: 100% (100%)
 18. 100% MATCH: 100% (100%)
 19. 100% MATCH: 100% (100%)
 20. 100% MATCH: 100% (100%)

Query Match: 40.00% Score: 40.00 DB: 4 Length: 112
 Best Local Similarity: 40.00% Prod. No. 1.40 40
 Matches: 100% (100%) Mismatches: 122% (100%) Gaps: 5%

01. 100% MATCH: 100% (100%)
 02. 100% MATCH: 100% (100%)
 03. 100% MATCH: 100% (100%)
 04. 100% MATCH: 100% (100%)
 05. 100% MATCH: 100% (100%)
 06. 100% MATCH: 100% (100%)
 07. 100% MATCH: 100% (100%)
 08. 100% MATCH: 100% (100%)
 09. 100% MATCH: 100% (100%)
 10. 100% MATCH: 100% (100%)
 11. 100% MATCH: 100% (100%)
 12. 100% MATCH: 100% (100%)
 13. 100% MATCH: 100% (100%)
 14. 100% MATCH: 100% (100%)
 15. 100% MATCH: 100% (100%)
 16. 100% MATCH: 100% (100%)
 17. 100% MATCH: 100% (100%)
 18. 100% MATCH: 100% (100%)
 19. 100% MATCH: 100% (100%)
 20. 100% MATCH: 100% (100%)

Query Match: 40.00% Score: 40.00 DB: 4 Length: 112
 Best Local Similarity: 40.00% Prod. No. 1.40 40
 Matches: 100% (100%) Mismatches: 100% (100%) Gaps: 5%

01. 100% MATCH: 100% (100%)
 02. 100% MATCH: 100% (100%)
 03. 100% MATCH: 100% (100%)
 04. 100% MATCH: 100% (100%)
 05. 100% MATCH: 100% (100%)
 06. 100% MATCH: 100% (100%)
 07. 100% MATCH: 100% (100%)
 08. 100% MATCH: 100% (100%)
 09. 100% MATCH: 100% (100%)
 10. 100% MATCH: 100% (100%)
 11. 100% MATCH: 100% (100%)
 12. 100% MATCH: 100% (100%)
 13. 100% MATCH: 100% (100%)
 14. 100% MATCH: 100% (100%)
 15. 100% MATCH: 100% (100%)
 16. 100% MATCH: 100% (100%)
 17. 100% MATCH: 100% (100%)
 18. 100% MATCH: 100% (100%)
 19. 100% MATCH: 100% (100%)
 20. 100% MATCH: 100% (100%)

Query Match: 40.00% Score: 40.00 DB: 4 Length: 112
 Best Local Similarity: 40.00% Prod. No. 1.40 40
 Matches: 100% (100%) Mismatches: 100% (100%) Gaps: 5%

01. 100% MATCH: 100% (100%)
 02. 100% MATCH: 100% (100%)
 03. 100% MATCH: 100% (100%)
 04. 100% MATCH: 100% (100%)
 05. 100% MATCH: 100% (100%)
 06. 100% MATCH: 100% (100%)
 07. 100% MATCH: 100% (100%)
 08. 100% MATCH: 100% (100%)
 09. 100% MATCH: 100% (100%)
 10. 100% MATCH: 100% (100%)
 11. 100% MATCH: 100% (100%)
 12. 100% MATCH: 100% (100%)
 13. 100% MATCH: 100% (100%)
 14. 100% MATCH: 100% (100%)
 15. 100% MATCH: 100% (100%)
 16. 100% MATCH: 100% (100%)
 17. 100% MATCH: 100% (100%)
 18. 100% MATCH: 100% (100%)
 19. 100% MATCH: 100% (100%)
 20. 100% MATCH: 100% (100%)

Query Match: 40.00% Score: 40.00 DB: 4 Length: 112
 Best Local Similarity: 40.00% Prod. No. 1.40 40
 Matches: 100% (100%) Mismatches: 100% (100%) Gaps: 5%


```

Query Match      28.28: Score 451: DB 11: Length 245:
Best Local Similarity 40.68: Pred. No. 4,2e-30:
Matches 99: Conservative 54: Mismatches 81: Indels 10: Gaps 4:

QY 1 7 6621SFANSTLLAAVSPTLLWLLINWSTVLNPAENVEVET ANIMAVINHEPS 95
   1 111 1 1111111 1111111 111 111 111 111 111 111 111 111
DB 1 6 6081SVVSTLALATSLIALMLLVSWMLPYQDMVLRKVSIMSTWITFNSS 65
   1 111111 111111 111111 111 111 111 111 111 111 111 111

QY 96 NMATTSITVLIKIANPSEIFLILKREKSVILVMLLPLEFLACHLFVINMNEIVRT 155
   1 111111 111111 111111 111 111 111 111 111 111 111 111
DB 66 1WPAESI STEVEYIAMSPTPIYVVPV KYVMIGETIMSTIIPY NLIEMATEPIL 125
   1 111111 111111 111111 111 111 111 111 111 111 111 111

QY 156 FEEDRMIMWLEESAMV FEEMIVMANLVLPILLSIMLHSSLSMILKMO 210
   1 111111 111111 111111 111 111 111 111 111 111 111 111
DB 126 TEVYVSMSTSLNRTQSLNIFPAN--IMF GEIFAVSLVFLVSLMRHQRMO 181
   1 111111 111111 111111 111 111 111 111 111 111 111 111

QY 211 LHRGSGIPISTKVIKAIQIVISPLALCAIYPLSIMSVSPGSLNKPVEPKAKIRPS 270
   1 111111 111111 111111 111 111 111 111 111 111 111 111
DB 182 HSAHSGPIASTKAIIPAIQITLIALSLVSTFPLSHVKKVSAIIEPTILHIVVAPTA 241
   1 111111 111111 111111 111 111 111 111 111 111 111 111

QY 271 YPSI 274
   111
DB 242 FPSV 245

```

Search completed: August 12, 2002, 13:57:47
 Job time: 458 sec



TITLE Direct Submission
JOURNAL Submitted (21-Dec-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On May 26, 2000 this sequence version replaced g1.6759210. All repeats were identified using RepeatMasker.
Shitt, A.F., A. S. Green, P. (1996-1997)
<http://http://genome.washington.edu/kcg/bioinformatics/bmi>
----- Genome Center -----
----- Genome Center for Genome Research -----

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIKR
 Web site: <http://www.sscg.wi.mit.edu>
 Contact: sequence.submissions@genome.wi.mit.edu
 Project Information
 Center project name: l3966
 Center clone name: l14.0.21
 Summary Statistics
 WIKR785: 100% of reads

Sequencing vector:	M13; M77815; 100% of reads
Chemistry:	Dye-terminator Big Dye [®] ; 100% of reads
Assembly program:	Phrap; version 0.960741
Consensus quality:	156273 bases at least Q40
Consensus quality:	159226 bases at least Q30
Consensus quality:	160285 bases at least Q20
Insert size:	167000; adaro5e-1p
Insert size:	161580; sum-of-confids
Quality coverage:	4.7 in Q20 bases; adaro5e-1p
Quality coverage:	4.8 in Q20 bases; sum-of-confids

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

3488: contig of 3488 bp in length

* be preserved. 3488: contig of 3488 bp in length
 * 1 3489 3588: gap of 100 bp
 * 3899 15445: contig of 11857 bp in length
 * 15446 15545: gap of 100 bp
 * 15546 27026: contig of 11481 bp in length
 * 27027 27126: gap of 100 bp
 * 27127 43966: contig of 16840 bp in length
 * 43967 44066: gap of 100 bp
 * 44067 60034: contig of 15968 bp in length
 * 60035 60134: gap of 100 bp
 * 60135 81964: contig of 21850 bp in length
 * 81965 82064: gap of 100 bp
 * 82065 103463: contig of 21359 bp in length
 * 103464 103563: gap of 100 bp
 * 103564 130052: contig of 26489 bp in length
 * 130053 130152: gap of 100 bp
 * 130153 162180: contig of 32228 bp in length.

```

FEATURES
source
location/Qualifiers
1. 162380
for/ganism: "Homo sapiens"
/db_xref: "taxon:9606"
/cclone: "R611.114021"
/cclone_id: "R6C1-11 Human Male RAC"
1. 3488
/mole: "assembly_11.treatment"
3689. 15345
/mole: "assembly_11.treatment"
15543. 27026
/mole: "assembly_11.treatment"
27127. 43066
/mole: "assembly_11.treatment"
clone_end: sp6
vector_side: right"
44067. 60034
/mole: "assembly_11.treatment"
60135. 81964
/mole: "assembly_11.treatment"

```

```
misc_feature      82065..103463  
                  /note="assembly_fragment"  
misc_feature      103562..136052  
                  /note="assembly_fragment"  
clone_end=17  
vector_side:left  
130153..162480  
/note="assembly_fragment"  
misc_feature      225722..242584  
BASH_COUNT       242584 g 49946 t  
ORIGIN
```

Query Match	88.58%	Score 822.8	DB 2	Length 162,880
Best Local Similarity	92.88%	Field No. 5, 2e-173		
Matches	863	Mismatches	67	Gaps 0

[illegible][illegible]

DB 124517
ATCTGCTTGGTCGACCAATTGTCACTGCTCTGGCGAGTCGCACAATTGTTTAACTCTGG 124

[illegible][illegible][illegible][illegible][illegible]

DB 124217 GCCTGTCACTCTTTGTGTGTAACACATGATCAGATTATATCGAG AAAACCAATATACAGGCG
 481 aacatgacattggaagatcgaattggaagatcgaattactcttccaaatattgaattatattc 54
 QY

541 atgtatgaagaatttaattaccttgaattctgacctatctttatgctttatctgt 600
Db 124157 AACATGACCTGGAGACATCAATTTGAGCGCTGCAAGTACCTTTCATATACATGATATATACCT 12

Db 124097 ATGCTAGCAACTTAACTACCTTTACTGTAAACCTGATATCTTCTGCTTAACTCTAT

DB 124037 TCTCTGTAAACATCTTCAGAACATCTAGCTCTATGGTAAAGGATCTTCAGACATCTTCACT 7

[illegible][illegible]

DB 123857 TTTCTGTTCTGCGCAAGCTATATGGATTCAGCTCTCTCTTAACTTACCTCTTCTATCTTCAAT

841 *ttggtgaatcgaagaacttadagcgaacttctcttcatttttgaatgaattgaattatctga*

repeat_rejoin	/rpt_family "AluIo"	complement (24608	24133)
repeat_rejoin	/rpt_family "GADn"	complement (28124	28214)
repeat_rejoin	/rpt_family "TAOn"	complement (28272	28381)
repeat_rejoin	/rpt_family "tIM4"	complement (28912	28968)
repeat_rejoin	/rpt_family "tIM4"	complement (29052	29102)
repeat_rejoin	/rpt_family "M11A1"	complement (29201	29267)
repeat_rejoin	/rpt_family "Alus"	complement (29508	29553)
repeat_rejoin	/rpt_family "GTAAOn"	complement (29560	29588)
repeat_rejoin	/rpt_family "AluO"	complement (30469	30498)
repeat_rejoin	/rpt_family "AT-rich"	complement (31316	31367)
repeat_rejoin	/rpt_family "t11A2"	complement (32070	32157)
repeat_rejoin	/rpt_family "AluIo"	complement (32684	33480)
repeat_rejoin	/rpt_family "t11A5"		
Score Match	74.38	Score 691.41	DB 92
Post-Repeat Similarity	84.68	Post. No. 7,30-144	
Matches	7897	Unscored 100	0
		Mismatches	141
		Indels	4
		Gaps	1
		Length	173735

[illegible][illegible]

RESULT 13																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
-----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

[illegible]

Search completed: August 12, 2002, 12:50:58
Job time: 655.1 sec



RP SEQUENCE FROM N.A.
 RX MEDLINE 2022571; PubMed-10761944;
 RA Adler E., Boon M.A., Mueller K.L., Chandrasekar J., Kyba N.J.P.,
 RA Zuker C.S.;
 RI "A novel family of mammalian taste receptors";
 RL J Biol Chem 274(20):10059-10064(2000);
 DR EMBL AF227142; AAF43415.1;
 KW Receptor;
 SO STUDENT 408 AA; 65052 MW; 00804194204E043 C8064;

Query Match 2.68; Score 8; DB 11; Length 408;
 Post Local Similarity 100.0%; Prod. No. 10;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 LKIANFSN 115
 128 LKIANFSN 115

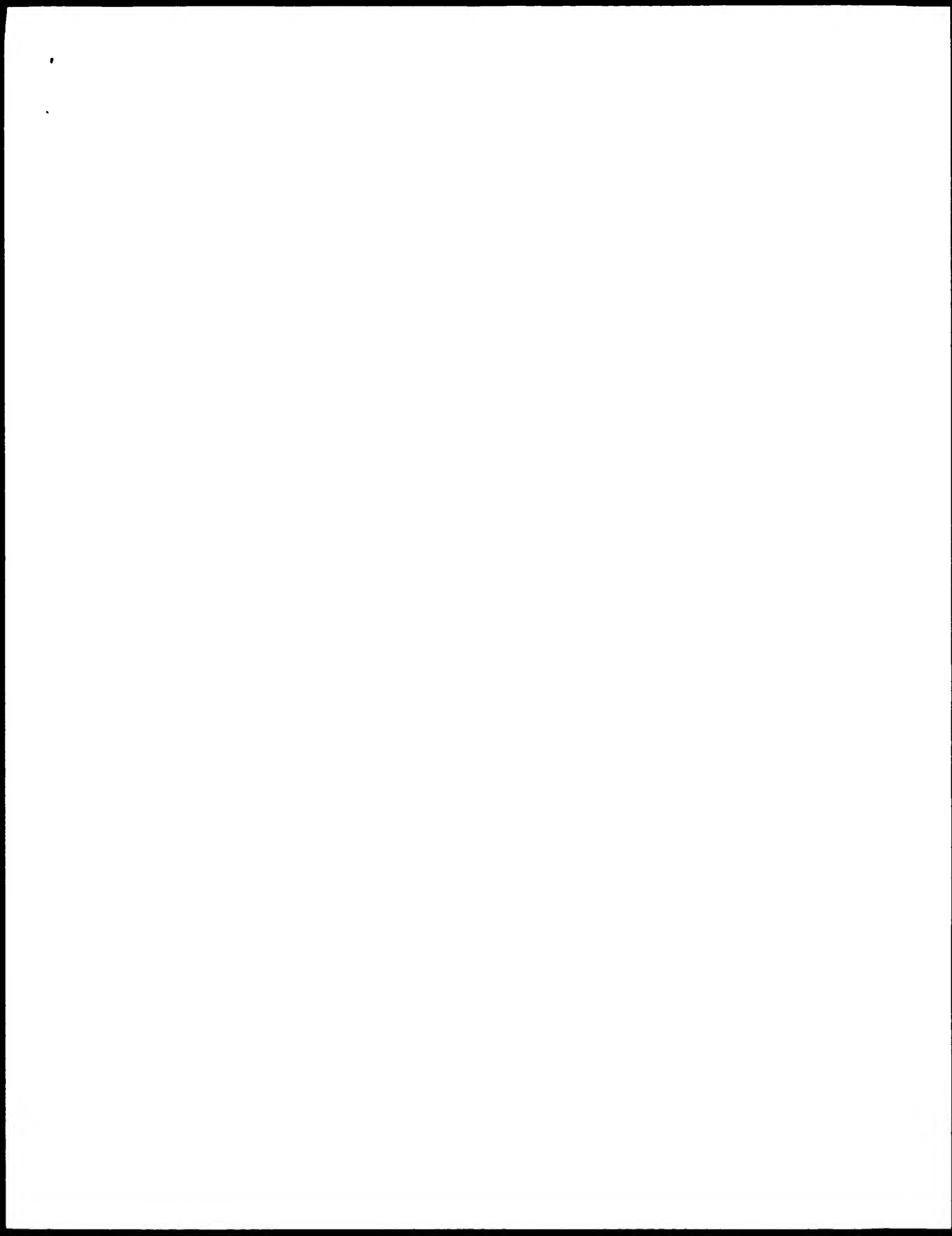
RESULT 15

Q96J93 PRELIMINARY; PRT; 411 AA;
 AC Q96J93;
 DE 01-JUN-2001 (FEBMRL17, created)
 DT 01-JUN-2001 (FEBMRL17, last sequence update)
 DI 01-OCT-2001 (FEBMRL18, last annotation update)
 DE GLYCOSYL TRANSFERASE;
 GN YBA1;
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis);
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococaceae;
 OX Lactococcus;
 OX NCBI_TaxID:1460;
 RN 11;
 RP SEQUENCE FROM N.A.
 RC STRAIN 121403;
 RX MEDLINE 21235186; PubMed-1143471;
 RA Bolotin A., Winkler P., Manger S., Jallou O., Malarme K.,
 RA Weisbach J., Ehrlich S.D., Sorokin A.;
 RI "The complete genome sequence of the lactic acid bacterium Lactococcus
 RI lactis ssp. lactis 121403";
 RL Genome Res. 11(7):754-764(2001);
 RK EMBL AF006249; AAK04055.1;
 DR InterPro:IPR001734; Glycosyl_Transf_2;
 DR Pfam:PF04545; Glycosyl_Transf_2;
 KW Transferase; Complete Proteome;
 SO SEQUENCE 411 AA; 35472 MW; 66089695384527 C8064;

Query Match 2.68; Score 8; DB 16; Length 411;
 Post Local Similarity 100.0%; Prod. No. 10;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 LSTFYLIK 159
 102 LSTFYLIK 159

Search completed: August 12, 2002, 14:00:49
 Job time: 266 sec



Version 4.5
Copyright (c) 1994 - 2000 CompuGen Ltd

ON PROBABILITIES IN SEARCHING MODELS

Run (ID): August 12, 2002, 14:51:14 ; Search time 24.7 seconds

484.386 Million cell updates/sec

See Table 10.2

105224 seqs, 48719550 residues

total number of hits satisfying chosen parameters: 105224

[illegible]

Maximum Likelihood Estimation

Post Processing: Minimum Match 08

Maximum Match 100%

Listed first 45 summaries

```
def abuse : SwissT(1_4) :=
```

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARY

Result	No.	Score	Match	Quarter	Month	Day	ID	Description
1	144	9.0	377	1	APR	RAT	Q973641 rat105 (not)	
2	147	8.6	377	1	APR	MOUSE	Q973638 mus musculi	
3	148.5	8.4	380	1	APR	MACAQUE	Q973664 maca muli	
4	150.5	8.2	380	1	APR	HUMAN	P35414 homo sapien[1]	
5	124	7.7	380	1	CHIK	HUMAN	P51679 homo sapien[2]	
6	124	7.7	453	1	CHIK	XENLA	P73041 xenopus laevis	
7	122	7.6	440	1	CAVE	STRAU	P30854 staphylococ	
8	121.5	7.6	345	1	CHIK	PIC	Q97300 sus scrofa	
9	121	7.6	388	1	SH4	MOUSE	P37288 mus musculi	
10	117	7.3	460	1	CHIK	MOUSE	P51640 mus musculi	
11	116	7.3	446	1	CHIK	HUMAN	Q96876 homo sapien[1]	
12	115.5	7.2	463	1	CHIK	HUMAN	P50052 homo sapien[2]	
13	115	7.2	358	1	CYH	HEISA	Q97042 hiliptephia	
14	114	7.1	488	1	SH4	HUMAN	Q13839 homo sapien[1]	
15	112.5	7.0	312	1	CHIK	RAT	P34987 rat105 (not)	
16	112	7.0	409	1	CHIK	RAT	Q972419 rat105 (not)	
17	112	7.0	480	1	CHIK	SHRIMP	P311203 scyllaridin	
18	111.5	7.0	463	1	CHIK	MOUSE	P35374 mus musculi	
19	111.5	7.0	463	1	CHIK	XENLA	P35373 xenopus laevis	
20	111	7.0	481	1	SH4	CAVE	Q970417 capri (pove)	
21	111	7.0	488	1	SH4	CAVE	Q970528 cavia porcu	
22	110.5	6.9	363	1	CHIK	MACAQUE	Q970476 macaques (not)	
23	110.5	6.9	463	1	CHIK	RAT	P35375 rat105 (not)	
24	110.5	6.9	533	1	CHIK	SHRIMP	P311203 scyllaridin	
25	109	6.8	409	1	CHIK	MOUSE	Q972041 mus musculi	
26	108.5	6.8	406	1	SH4	RAT	Q972758 rat105 (not)	
27	108	6.8	382	1	CHIK	MOUSE	Q97277 rat105 (not)	
28	107.5	6.7	346	1	SH4	CAVE	P345672 cavia porcu	
29	107.5	6.7	359	1	CHIK	CAVE	Q969926 capri (pove)	
30	106.5	6.7	352	1	CHIK	HUMAN	P340941 homo sapien[1]	
31	106.5	6.7	353	1	APR	XENLA	P373993 xenopus laevis	
32	106.5	6.7	359	1	CHIK	CHICK	P373993 gallus gall	
33	106.5	6.7	359	1	CHIK	MACAQUE	P343399 macaques (not)	

RESULT	1	ALIGNMENTS
ADJ_RAT	STANDARD:	PRG: 477 AA.
ADJ_RAT	Q9J0B3: Q9ESK2:	
D1	01-MAR-2002 (Ref. 41, created)	
D1	01-MAR-2002 (Ref. 41, last sequence update)	
D1	01-MAR-2002 (Ref. 41, last annotation update)	
D6	Apelin receptor (G protein coupled receptor APJ) (Angiotensin receptor-like 1) (B7b) (GPR44).	
D6	APJL1 (APJ).	
G8	Rattus norvegicus (Rat).	
G8	Elkayot; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
G8	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
G8	NCBI_Taxid 10116;	
K8	SEQUENCE FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 20447212; PubMed 10777510;	
KA	Kishida C., Kawamura Y., Fukushima S., Fujii R., Habata Y., Hinuma S.,	Q9J0B3: closest hit
KA	Kishida C., Hiyata S., Kurokawa T., Oda H., Nishimura Y., Fujita M.,	P43554: mus muscul
KT	"Molecular and functional characteristics of APJ: Tissue distribution of mRNA and interaction with the endogenous ligand apelin."	P43977: Rattus s. t
KT	J. Biol. Chem. 275:21061-21067(2000).	P40556: homo sapien
K8	SEQUENCE FROM N.A.	P41145: homo sapien
K8	TISSUE: Brain;	P43905: Chordatus c
K8	MEDLINE: 20447212; PubMed 10777510;	P43905: homo sapien
KA	Kishida C., Kawamura Y., Fukushima S., Fujii R., Habata Y., Hinuma S.,	P43905: sus scrofa
KA	Kishida C., Hiyata S., Kurokawa T., Oda H., Nishimura Y., Fujita M.,	P408707: mus muscul
KT	"Molecular and functional characteristics of APJ: Tissue distribution of mRNA and interaction with the endogenous ligand apelin."	G24474: mus musc
KT	J. Biol. Chem. 275:21061-21067(2000).	P41144: evula porci
K8	SEQUENCE FROM N.A.	G94972: Rattus s. t
K8	TISSUE: Brain;	
K8	MEDLINE: 20461761; PubMed 11004481;	
KA	O'Carroll A.-M., Selby J.L., Polkovic M., Tolatt S.J.,	
K7	"Identification of mRNA expression in rat adipose tissue, the rat adipin receptor, and its endogenous ligand apelin in brain and periperal tissues."	
KT	Proc. Natl. Acad. Sci. USA 97:1197-1201(2000).	
KT	Proc. Natl. Acad. Sci. USA 97:1197-1201(2000).	
K8	SEQUENCE FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	
K8	TISSUE: Brain;	
K8	MEDLINE: 21064694; PubMed 11144423;	
KA	De Mota N., Lelkes Z., Llorens Vives C.,	
KA	"Cloning, pharmacological characterization and brain distribution of the rat apelin receptor."	
KT	Neuroendocrinology 72:400-407(2000).	
K8	SEQUENCE OF 75 400 FROM N.A.	

FI CARBOHYD 174 173 N-LINKED (GLUCNA...) (POTENTIAL)
 SEQ SEQUENCE 477 AA: 42266 MM: AAA4A7509A2420B (K6764)

Query Match
 Post Local Similarity 19.48% Score 1372 DB 1: Length 4772
 Matches 663 Conserved 725 Mismatches 129 Indels 71 Gaps 14

6 LFSSIVVVEVTEVDEARFNAVNTW IKPFLTAA...
 28 LPAALYMLVLLGLTGNLTV... WIVKPSSEKRSADFLASLAAV...
 44 LKAVTV...
 80 LPLALTYRRTWTFSTKLSSTLFF... VNMASV...
 118 LKAVTV...
 136 ARKLAVAAVAIAVLAVALAVVAFSLASNTLQ...
 172 MWVTKKSAVPSNMTVAVN...
 194 WAVEVGL...
 222 FVHKAIVTEVTEVDEARFNAVNTW...
 248 LKAVTV...
 267 LKAVTV...
 295 LKAVTV...

RESULT 4
 AC1_MAMMOT STANDARD PRT: 480 AA
 AT 097664
 DT 16-JUN-2001 (Rel. 40, created)
 DT 16-JUN-2001 (Rel. 40, last sequence update)
 DT 01-MAR-2002 (Rel. 41, last annotation update)
 DE Apelin receptor (G protein coupled receptor A6) (Apoelousin
 DE receptor-like 1)
 GN AGRL1 or APD
 OS Macaca mulatta (Rhesus macaque)
 OS Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
 OS Mammalia: Eutheria: Primates: Catarrhini: Cercopithecoidea:
 OS Cercopithecoidea: Macaca
 NCBI_TaxId=9544
 RN [1]
 RP SEQUENCE FROM N.A.
 RC 11SSDE-SPEECH
 RA MEDLINE 21354176: Pubmed 11461684
 RA Margulies H.J., Hauer D.A., Clements J.E.
 RT Identification and comparison of eleven Rhesus macaque chemokine
 RT receptors.
 BL AIDS Res. Hum. Retroviruses 17:981-986(2001)
 FT -1- FUNCTION: Receptor for apelin coupled to G proteins that inhibit
 FT adenylyl cyclase activity. Alternative coreceptor with CPM for
 FT HIV infection.
 FT SUBCELLULAR LOCATION: Integral membrane protein.
 FT SIMILARITY: BELONGS TO FAMILY 1 OF G PROTEIN-COUPLED RECEPTORS.

This SWISS-Prot entry is copyright. It is produced through a collaboration
 between The Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/submit/submit.html>
 or send an email to license@isb.scrib.ch).

EMBL: AF100236; AA72404.1
 RefSeq: PR000276; cPCR_Throspen
 PIR: P00001; Tm_1_1

DR PRINTS: PR00237; cPCR_Throspen
 DR PROSITE: PS00237; G PROTEIN RECEPTOR_11_1
 DR PROSITE: PS00237; G PROTEIN RECEPTOR_11_1
 KW G protein coupled receptor, transmembrane, glycoprotein
 FT DOMAIN 1 26
 FT TRANSMM 27 51
 FT TRANSMM 52 66
 FT TRANSMM 67 91
 FT TRANSMM 92 125
 FT TRANSMM 126 144
 FT TRANSMM 145 166
 FT TRANSMM 167 201
 FT TRANSMM 202 221
 FT TRANSMM 222 241
 FT TRANSMM 242 271
 FT TRANSMM 272 284
 FT TRANSMM 285 308
 FT TRANSMM 309 330
 FT TRANSMM 331 350
 FT TRANSMM 351 375
 FT TRANSMM 376 395
 FT TRANSMM 396 426
 FT TRANSMM 427 446
 FT TRANSMM 447 466
 FT TRANSMM 467 480
 FT TRANSMM 481 500
 FT TRANSMM 501 520
 FT TRANSMM 521 540
 FT TRANSMM 541 560
 FT TRANSMM 561 580
 FT TRANSMM 581 600
 FT TRANSMM 601 620
 FT TRANSMM 621 640
 FT TRANSMM 641 660
 FT TRANSMM 661 680
 FT TRANSMM 681 700
 FT TRANSMM 701 720
 FT TRANSMM 721 740
 FT TRANSMM 741 760
 FT TRANSMM 761 780
 FT TRANSMM 781 800
 FT TRANSMM 801 820
 FT TRANSMM 821 840
 FT TRANSMM 841 860
 FT TRANSMM 861 880
 FT TRANSMM 881 900
 FT TRANSMM 901 920
 FT TRANSMM 921 940
 FT TRANSMM 941 960
 FT TRANSMM 961 980
 FT TRANSMM 981 1000

Query Match
 Post Local Similarity 18.78% Score 1345 DB 1: Length 4800
 Matches 663 Conserved 725 Mismatches 129 Indels 71 Gaps 14

6 LFSSIVVVEVTEVDEARFNAVNTW IKPFLTAA...
 28 LPAALYMLVLLGLTGNLTV... WIVKPSSEKRSADFLASLAAV...
 44 LKAVTV...
 80 LPLALTYRRTWTFSTKLSSTLFF... VNMASV...
 118 LKAVTV...
 136 ARKLAVAAVAIAVLAVALAVVAFSLASNTLQ...
 172 MWVTKKSAVPSNMTVAVN...
 194 WAVEVGL...
 222 FVHKAIVTEVTEVDEARFNAVNTW...
 248 LKAVTV...
 267 LKAVTV...
 295 LKAVTV...

RESULT 4
 AC1_HUMAN STANDARD PRT: 480 AA
 AT 097664
 DT 01-JUN-1994 (Rel. 29, created)
 DT 01-JUN-1994 (Rel. 29, last sequence update)
 DT 01-MAR-2002 (Rel. 41, last annotation update)
 DE Apelin receptor (G protein coupled receptor A6) (Apoelousin
 DE receptor-like 1) (db11)
 GN AGRL1 or APD
 OS Homo sapiens (human)
 OS Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
 OS Mammalia: Eutheria: Primates: Catarrhini: Hominidae: Homo
 NCBI_TaxId=9606
 RN [1]
 RP SEQUENCE FROM N.A.
 RC 11SSDE-SPEECH
 RA MEDLINE 94124041: Pubmed 8240432
 RA Kennedy J.L., Shi X., Polonais A., George S.P., Nayyan I.
 RA Kennedy J.L., Shi X., Polonais A., George S.P., Nayyan I.
 RT A human gene that shows identity with the gene encoding the

EMBL: AF100236; AA72404.1
 RefSeq: PR000276; cPCR_Throspen
 PIR: P00001; Tm_1_1

[illegible][illegible]

[illegible]

• • •

Revised version 4.5
Copyright (c) 1994 - 2000 Compucon Ltd

IMPROVING PROTEIN QUALITY USING A MODEL

Run on: August 12, 2002, 13:58:38 ; Search time 14.59 seconds

(without alijments)

115-19-M25-882-8

Sequence: 1 MTFPI PFSI VVPIVIG FLSVFQMBRYWKZHTSSP 409

Summary table:

Searched: 105224 seqs, 3871950 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum leaf length: (1)

Maximum job seq length: 2000000000

Post-processing: Listing first 45 summaries

100 + 20 + 4 = 124

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARY

Fossil No.	Score	Match	Length	DB	ID	Insertion
1	8	2.6	484	1	MRCR_MOUSE	p57410 bartonia ap
2	7	2.3	83	1	YXCA_BAC50	p46133 bacillus su
3	7	2.3	131	1	KV41_MOUSE	p01661 mus musculu
4	7	2.3	132	1	RV4F_MOUSE	p01661 mus musculu
5	7	2.3	164	1	MNRB_CHIMP	p04188 chimpanza m
6	7	2.3	284	1	APX1_STRCR	p08561 streptomyces
7	7	2.3	286	1	FINB_ECOLI	p08186 escherichia
8	7	2.3	320	1	PSST_PANST	p08116 escherichia
9	7	2.3	327	1	PS12_RHLA	p08116 escherichia
10	7	2.3	437	1	GRTE_ECOLI	p08155 escherichia
11	7	2.3	439	1	LINE_ECOLI	p09444 escherichia
12	7	2.3	492	1	CP01_PPMV	p09444 escherichia
13	7	2.3	552	1	NKPE_ECOLI	p42716 escherichia
14	7	2.3	555	1	Y143_YEAST	p06115 saccharomyce
15	7	2.3	568	1	NBM2_MOUSE	p49281 mus musculu
16	7	2.3	568	1	NBM2_PAT	p49281 mus musculu
17	7	2.3	576	1	W376_YDAST	p50906 sporothrix
18	7	2.3	584	1	NIFA_RHET	p54941 rhizobium c
19	7	2.3	587	1	RHC2_RAT	p70545 rat ius corti
20	7	2.3	615	1	R1A1_AZOCA	p09114 azotobacter
21	7	2.3	635	1	NKPE_HAETN	p44944 haemophilus
22	7	2.3	686	1	RECV_STAUB	p050581 staphylococ
23	7	2.3	862	1	ABHE_GLOAB	p34744 clostridium
24	7	2.3	896	1	PR0P_NKPEP	p34540 pneumocyst
25	4	1.9	24	1	TEP4_L471B	p060231 thioridol
26	6	1.9	57	1	SK11_MESMA	p06116 bacillus mus
27	6	1.9	58	1	SPF2_MESMA	p06116 bacillus mus
28	6	1.9	65	1	VY13_VACCIN	p21044 vaccinia v
29	6	1.9	76	1	ALP9_HANWI	p48881 haemophilus
30	6	1.9	77	1	YGRG_RACE1	p46132 haemophilus
31	6	1.9	82	1	YGRD_RHCPD	p04919 bacteriophage
32	6	1.9	84	1	RGH1_HUMAN	p06076 hamo sapien
33	6	1.9	86	1	TXW1_NAVAL	p09412 naja atra

ALIGNMENTS

45	6	1.9	87	1	NX1.3-0.1NMJ
44	6	1.9	97	1	NX1.6-0.0NMJ
46	6	1.9	87	1	VAL.2-PA07.1
37	6	1.9	89	1	CY1.3-0.0ANA
38	6	1.9	102	1	UJ7.3-FEV
49	6	1.9	104	1	RS1.49-0.1MA
40	6	1.9	106	1	RS10.1-MY-FEV
41	6	1.9	115	1	NO10.1-0.0MA
42	6	1.17	1	NO10.1-0.1MA	
43	6	1.9	317	1	CX2.2-CY1.0FE
44	6	1.9	120	1	GLH1-PA1
45	6	1.9	120	1	GLH2-PA1
46	6	1.9	120	1	GLH2-PA1

RESULT	1
MURC_BUCAL	
1D	
MURC_BUCAL	
STANDARD;	
PRT;	484 AA

107 16-OCF-2001 (Ref., 40), (created)
 108 16-OCF-2001 (Ref., 40), last sequence update
 109 16-OCF-2001 (Ref., 40), last annotation update
 110 UDP-N-acetylglutamate:alanine ligase (EC 6.3.2.8) (UDP-N
 111 acetylglutamate-L-alanine synthetase).

08 bacterium aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
09 symbol to bacterium).
10 Rectal prolapse; rectocele; hemorrhoids.
11 NBI_TaxID_116099.

16. STEININ TOKYO 1998;
MEDLINE:20445173; PubMed:1394077;
FA Shigenobu S., Matsuda H., Hattori M., Sekiya Y., Ishikawa H. et al.
"Genome sequence of the endocellular bacterial symbiont of aphids

62	-1- FUNCTION: CELL WALL FORMATION (BY SIMILARITY).	
63	-1- CATALYTIC ACTIVITY: ADP + (phospho)-acetylthioladenyl + L-alanine	ADP
64	phosphate + (phospho)-acetylthioladenyl + L-alanine.	
65	-1- PATHWAY: PEPTIDOLYAN BIOSYNTHESIS.	

This SWI-Protein entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the Max-Planck-Gesellschaft. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is not in any way modified and this statement is not removed, altered by and for commercial purposes (e.g. for advertising or promotional purposes, for creating new collective works, or for resale) or send an email to license@ebi.ac.uk.

Lac	P130M; P90(125); Mut. frequency: 1.
Km	Diploid yeast synthetic; cod + wld + col + histidin ⁻ ; Lysase
Km	AIP-binding; Complete proteome.
RN	RN-B10d 125 141 AIP (POTENTIAL).
Su	STRUCTURE 484 AA; 5411 MB; 06-85/04-57/55/90/07-Br/64;

Query Match	2.68	Score 8	106.17	Length 484
Post Local Similarity	100.00	Prod. No.	4.27	
Matches	8	Conservative	0	Mismatches
			0	Indels
			0	Gaps

RESULT AND YXCA_FACTSU


```

01  Bacteriophages: alpha subdivision; Rhizobiales group;
02  Plectonictes: Mesorhizobium.
03  Plectonictes: Mesorhizobium.
04  Mesorhizobium.
05  Mesorhizobium.
06  Mesorhizobium.
07  Mesorhizobium.
08  Mesorhizobium.
09  Mesorhizobium.
10  Mesorhizobium.
11  Mesorhizobium.
12  Mesorhizobium.
13  Mesorhizobium.
14  Mesorhizobium.
15  Mesorhizobium.
16  Mesorhizobium.
17  Mesorhizobium.
18  Mesorhizobium.
19  Mesorhizobium.
20  Mesorhizobium.
21  Mesorhizobium.
22  Mesorhizobium.
23  Mesorhizobium.
24  Mesorhizobium.
25  Mesorhizobium.
26  Mesorhizobium.
27  Mesorhizobium.
28  Mesorhizobium.
29  Mesorhizobium.
30  Mesorhizobium.
31  Mesorhizobium.
32  Mesorhizobium.
33  Mesorhizobium.
34  Mesorhizobium.
35  Mesorhizobium.
36  Mesorhizobium.
37  Mesorhizobium.
38  Mesorhizobium.
39  Mesorhizobium.
40  Mesorhizobium.
41  Mesorhizobium.
42  Mesorhizobium.
43  Mesorhizobium.
44  Mesorhizobium.
45  Mesorhizobium.
46  Mesorhizobium.
47  Mesorhizobium.
48  Mesorhizobium.
49  Mesorhizobium.
50  Mesorhizobium.
51  Mesorhizobium.
52  Mesorhizobium.
53  Mesorhizobium.
54  Mesorhizobium.
55  Mesorhizobium.
56  Mesorhizobium.
57  Mesorhizobium.
58  Mesorhizobium.
59  Mesorhizobium.
60  Mesorhizobium.
61  Mesorhizobium.
62  Mesorhizobium.
63  Mesorhizobium.
64  Mesorhizobium.
65  Mesorhizobium.
66  Mesorhizobium.
67  Mesorhizobium.
68  Mesorhizobium.
69  Mesorhizobium.
70  Mesorhizobium.
71  Mesorhizobium.
72  Mesorhizobium.
73  Mesorhizobium.
74  Mesorhizobium.
75  Mesorhizobium.
76  Mesorhizobium.
77  Mesorhizobium.
78  Mesorhizobium.
79  Mesorhizobium.
80  Mesorhizobium.
81  Mesorhizobium.
82  Mesorhizobium.
83  Mesorhizobium.
84  Mesorhizobium.
85  Mesorhizobium.
86  Mesorhizobium.
87  Mesorhizobium.
88  Mesorhizobium.
89  Mesorhizobium.
90  Mesorhizobium.
91  Mesorhizobium.
92  Mesorhizobium.
93  Mesorhizobium.
94  Mesorhizobium.
95  Mesorhizobium.
96  Mesorhizobium.
97  Mesorhizobium.
98  Mesorhizobium.
99  Mesorhizobium.
100  Mesorhizobium.

```


Query Match 2.000 Score 72.000 Length 200
 Best Local Similarity 100.000 Prod. No. 4.62
 Matches 71 Conservative 02 Mismatches 02 Indels 02 Gaps 02

QY 58 ELWELL 64
 DB 7 ELWELL 14

RESULT 2
 US-09-095-659-16
 Sequence 16, Application US/080905659
 Patent No. 6,242,213
 GENERAL INFORMATION:
 APPLICANT: Anderson, Dirk M.
 APPLICANT: Marosovsky, Eugene
 APPLICANT: Lauroit, Gilbert
 TITLE OF INVENTION: Method for receptor Activation of NF-kappaB
 NUMBER OF SEQUENCES: 19
 CORRESPONDENT ADDRESS:
 ADDRESSEE: Immunovir Corporation, Law Department
 STREET: 11 University Street
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZITE: 0001
 COMPUTER REAMABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: Apple Power Macintosh
 OPERATING SYSTEM: Apple operating System 7.5.5
 SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
 CURRENT APPLICATION DATA:
 FILING DATE: 22 DECEMBER 1997
 CLASSIFICATION:
 APPLICATION NUMBER: US97 602964 671
 FILING DATE: 14 OCTOBER 1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 FILING DATE: 07 MARCH 1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 FILING DATE: 23 DECEMBER 1996
 CLASSIFICATION:
 ALTERNATIVE/AGENT INFORMATION:
 NAME: Lockhart, Patricia Anne
 REGISTRATION NUMBER: 44,693
 REFERENCE/WORK NUMBER: 2052 A
 INFORMATION INFORMATION:
 TELEPHONE: (206) 433-0644
 TELEFAX: (206) 433-0644
 INFORMATION FOR SEQ. ID NO. 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 MOLECULE TYPE: protein
 US-09-095-659-16

Query Match 2.000 Score 72.000 Length 200
 Best Local Similarity 100.000 Prod. No. 4.62
 Matches 71 Conservative 02 Mismatches 02 Indels 02 Gaps 02

QY 58 ELWELL 64
 DB 7 ELWELL 14

RESULT 3
 US-09-215-649A-16
 Sequence 16, Application US/09215649A
 Patent No. 6,271,849
 GENERAL INFORMATION:
 APPLICANT: Anderson, Dirk M.
 APPLICANT: Marosovsky, Eugene
 APPLICANT: Lauroit, Gilbert
 TITLE OF INVENTION: Receptor Activation of NF-kappaB
 NUMBER OF SEQUENCES: 19
 CORRESPONDENT ADDRESS:
 ADDRESSEE: Immunovir Corporation, Law Department
 STREET: 11 University Street
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZITE: 00101
 COMPUTER REAMABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: Apple Power Macintosh
 OPERATING SYSTEM: Apple operating System 7.5.5
 SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
 CURRENT APPLICATION DATA:
 FILING DATE: 17 JUNE 1998
 CLASSIFICATION: unknown
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/996,149
 FILING DATE: unknown
 APPLICATION NUMBER: US97 602964 671
 FILING DATE: 14 OCTOBER 1997
 CLASSIFICATION:
 APPLICATION NUMBER: US97 602964 671
 FILING DATE: 14 OCTOBER 1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 FILING DATE: 07 MARCH 1997
 CLASSIFICATION:
 APPLICATION NUMBER: US97 602964 671
 FILING DATE: 14 OCTOBER 1997
 CLASSIFICATION:
 ALTERNATIVE/AGENT INFORMATION:
 NAME: Lockhart, Patricia Anne
 REGISTRATION NUMBER: 44,693
 REFERENCE/WORK NUMBER: 2051-A
 INFORMATION INFORMATION:
 TELEPHONE: (206) 587-0430
 TELEFAX: (206) 243-0644
 INFORMATION FOR SEQ. ID NO. 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 MOLECULE TYPE: protein
 US-09-215-649A-16

Query Match 2.000 Score 72.000 Length 200
 Best Local Similarity 100.000 Prod. No. 4.62
 Matches 71 Conservative 02 Mismatches 02 Indels 02 Gaps 02

QY 58 ELWELL 64
 DB 7 ELWELL 14

RESULT 4
 US-09-320-424-25
 Sequence 25, Application US/09320424
 Patent No. 6,284,246
 GENERAL INFORMATION:
 APPLICANT: Wiloy, Steven R.
 APPLICANT: Goodman, Raymond G.
 TITLE OF INVENTION: Cytokine that Induces Apoptosis
 FILE REFERENCES: 2846 B
 CURRENT APPLICATION DATA:
 FILING DATE: 1999 05 26
 EARLIER APPLICATION NUMBER: 09/190,346
 EARLIER FILING DATE: 1998 11 10
 EARLIER APPLICATION NUMBER: 09/048,041

1 EARLIER FILING DATE: 1998-03-26
 2 EARLIER APPLICATION NUMBER: 08/670,454
 3 EARLIER FILING DATE: 1996-06-25
 4 EARLIER APPLICATION NUMBER: 08/548,468
 5 EARLIER FILING DATE: 1995-11-01
 6 EARLIER APPLICATION NUMBER: 08/496,632
 7 EARLIER FILING DATE: 1995-06-29
 8 NUMBER OF SEQ. ID NOS: 25
 9 SOFTWARE: Patent In Ver. 2.0
 10 SEQ. ID NO: 25
 11 LENGTH: 20
 12 TYPE: PRT
 13 ORGANISM: Homo sapiens
 14 US-09-320-424-25

Query Match 2.88: Score 7: DB 4: Length 20:
 Best Local Similarity 100.0% Prod. No. 4.62
 Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 58 ILWILL 64
 DB 7 ILWILL 13

RESULT 5
 US-08-518-845-5
 1 Sequence 5: Application US/08/518845
 2 Patent No. 6017754
 3 GENERAL INFORMATION:
 4 APPLICANT: CHRISTOPHER JONATHAN D
 5 APPLICANT: HEEFER, JAMES P.
 6 TITLE OF INVENTION: ROSE SYSTEM FOR ISOLATING AND
 7 TITLE OF INVENTION: IDENTIFYING POKERFISH CELLS TRANSFECTED WITH GENES AND
 8 NUMBER OF INVENTIONS: 2
 9 CORRESPONDENCE ADDRESS:
 10 ADDRESS: SOCIETY, SCOTT, MORRIS & PENSER
 11 STREET: 400 GARDEN CITY PLAZA
 12 CITY: GARDEN CITY
 13 STATE: NEW YORK
 14 COUNTRY: USA
 15 ZIP: 11530
 16 COMPUTER READABLE FORM:
 17 MEDIUM TYPE: floppy disk
 18 COMPUTER: IBM PC compatible
 19 OPERATING SYSTEM: PC-DOS/MS-DOS
 20 SOFTWARE: Patent In Release #1.0, Version #1.30
 21 CURRENT APPLICATION DATA:
 22 APPLICATION NUMBER: US/08/518,835
 23 FILING DATE: 24-06-1995
 24 CLASSIFICATION: 435
 25 ATTORNEY/AGENT INFORMATION:
 26 NAME: DIGILLO, FRANK S.
 27 REGISTRATION NUMBER: 41,446
 28 REFERENCE/SEQUENCE NUMBER: 11094
 29 TELECOMMUNICATION INFORMATION:
 30 TELEPHONE: (516) 742-4443
 31 TELEFAX: (516) 742-4366
 32 INFORMATION FOR SEQ. ID NOS: 5:
 33 SEQUENCE CHARACTERISTICS:
 34 LENGTH: 21 amino acids
 35 TYPE: amino acid
 36 STRANDEDNESS: single
 37 TOPOLOGY: linear
 38 US-08-518-845-5

Query Match 2.88: Score 7: DB 3: Length 21:
 Best Local Similarity 100.0% Prod. No. 4.62
 Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 QY 58 ILWILL 64

DB 7 ILWILL 13

RESULT 6
 US-07-644-278-21
 1 Sequence 21: Application US/07/644278
 2 Patent No. 5580101
 3 GENERAL INFORMATION:
 4 APPLICANT: COHEN, GARY L.
 5 APPLICANT: CO, Man Sung
 6 APPLICANT: SCHNEIDER, William P.
 7 APPLICANT: LAMOLFI, Nicholas P.
 8 APPLICANT: COLLING, Kathleen L.
 9 APPLICANT: SELICK, Harold E.
 10 TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
 11 NUMBER OF SEQUENCES: 11
 12 CORRESPONDENCE ADDRESS:
 13 ADDRESS: Townsend and Townsend Kiorito and Crow
 14 STREET: 379 Lytton Avenue
 15 CITY: Palo Alto
 16 STATE: California
 17 COUNTRY: US
 18 ZIP: 94301
 19 COMPUTER READABLE FORM:
 20 MEDIUM TYPE: floppy disk
 21 COMPUTER: IBM PC compatible
 22 OPERATING SYSTEM: PC-DOS/MS-DOS
 23 SOFTWARE: Patent In Release #1.0, Version #1.25
 24 CURRENT APPLICATION DATA:
 25 APPLICATION NUMBER: US/07/644,278
 26 FILING DATE: 19-DEC-1990
 27 CLASSIFICATION: 424
 28 ERROR APPLICATION DATA:
 29 APPLICATION NUMBER: US 07/590,274
 30 FILING DATE: 28 SEP-1990
 31 PRIOR APPLICATION DATA:
 32 APPLICATION NUMBER: US 07/310,252
 33 FILING DATE: 13-FEB-1989
 34 PRIOR APPLICATION DATA:
 35 APPLICATION NUMBER: US 07/290,975
 36 FILING DATE: 28-06-1988
 37 ATTORNEY/AGENT INFORMATION:
 38 NAME: SMITH, William M
 39 REGISTRATION NUMBER: 30,223
 40 REFERENCE/SEQUENCE NUMBER: 11823-002600
 41 TELECOMMUNICATION INFORMATION:
 42 TELEPHONE: (415) 326-2420
 43 INFORMATION FOR SEQ. ID NOS: 21:
 44 SEQUENCE CHARACTERISTICS:
 45 LENGTH: 126 amino acids
 46 TYPE: amino acid
 47 TOPOLOGY: linear
 48 MOLECULE TYPE: protein
 49 US 07 644 278 21

Query Match 2.88: Score 7: DB 1: Length 126:
 Best Local Similarity 100.0% Prod. No. 24:
 Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 QY 58 ILWILL 64
 DB 7 ILWILL 13

RESULT 7
 US-08-477-728-21
 1 Sequence 21: Application US/08/477728
 2 Patent No. 5585089
 3 GENERAL INFORMATION:
 4 APPLICANT: QUEEN, Gary L.

```

1 APPLICANT: SCHNEIDER, WILLIAM P.
2 APPLICANT: SELLER, HAROLD E.
3 TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
4 NUMBER OF SEQUENCES: 114
5 CORRESPONDENCE ADDRESS:
6 ADDRESSEE: Townsend and Townsend and Crew LLP
7 STREET: Two International Center, 8th Floor
8 CITY: Palo Alto
9 STATE: California
10 COUNTRY: US
11 ZIP: 94301
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Floppy disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: pc dos/MS-DOS
17 SOFTWARE: Patcutin Release #1.0, Version #1.25
18
19 PUBLICATION DATE: 02/07/95
20 FILING DATE: 07/09/90
21 PRIORITY DATE: 07/09/90
22 APPLICATION NUMBER: US 07/410,252
23 PUBLICATION NUMBER: 11825-002609
24
25 PRIOR APPLICATION DATA:
26 FILING DATE: 19/08/1990
27 APPLICATION NUMBER: US 07/290,975
28 FILING DATE: 28/08/1988
29 APPLICATION NUMBER: 28 SEP 1990
30 NAME: SMITH, WILLIAM M.
31 REGISTRATION NUMBER: 40,224
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: (415) 426-2400
34 TELEFAX: (415) 426-2422
35 INFORMATION FOR SEQ ID NO: 21:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 126 amino acids
38 TYPE: amino acid
39 TOPOLOGY: linear
40 MOLECULE TYPE: protein
41 US-08-474-040-21
42
43 Query Match: 2.00, Score 72, DB 1, Length 126;
44 Post Local Similarity: 100.00; Prod. No. 24;
45 Matches: 7; Conservation: 0; Mismatches: 0; Indels: 0;
46 Gaps: 0;
47
48 QY: 58 LWILL: 64
49 ID: 7 LWILL: 14
50
51 RESULT:
52 US-08-474-040-21
53 Sequence 21, April 04/00 US/08474040
54 Patent No. 569372
55 GENERAL INFORMATION:
56 APPLICANT: GILLEN, Gary L.
57 APPLICANT: GILLEN, Gary L.
58 APPLICANT: SCHNEIDER, William P.
59 APPLICANT: LANE, Nicholas P.
60 APPLICANT: GILLEN, Kathleen E.
61 APPLICANT: SELLER, Harold E.
62 TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
63 NUMBER OF SEQUENCES: 114
64 CORRESPONDENCE ADDRESS:
65 ADDRESSEE: Townsend and Townsend and Crew
66 STREET: 879 Lydon Avenue
67 CITY: Palo Alto
68 STATE: California
69 COUNTRY: US
70 ZIP: 94301
71 COMPUTER READABLE FORM:
72 MEDIUM TYPE: Floppy disk
73 COMPUTER: IBM PC compatible
74 OPERATING SYSTEM: pc dos/MS-DOS
75

```

```

1 CITY: Palo Alto
2 STATE: California
3 COUNTRY: US
4 ZIP: 94301
5
6 COMPUTER READABLE FORM:
7 MEDIUM TYPE: Floppy disk
8 COMPUTER: IBM PC compatible
9 OPERATING SYSTEM: pc dos/MS-DOS
10 SOFTWARE: Patcutin Release #1.0, Version #1.25
11
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US-09-474,040
14 FILING DATE: 07-09-1995
15 CLASSIFICATION: 546
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: US 07/410,252
18 FILING DATE: 19/08/1990
19 APPLICATION NUMBER: US 07/590,274
20 FILING DATE: 28-SEP-1990
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US 07/290,975
23 FILING DATE: 13/08/1988
24 APPLICATION NUMBER: 28 SEP 1990
25 NAME: SMITH, WILLIAM M.
26 REGISTRATION NUMBER: 40,224
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: (415) 426-2400
29 TELEFAX: (415) 426-2422
30 INFORMATION FOR SEQ ID NO: 21:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 126 amino acids
33 TYPE: amino acid
34 TOPOLOGY: linear
35 MOLECULE TYPE: protein
36 US-08-474-040-21
37
38 Query Match: 2.00, Score 72, DB 1, Length 126;
39 Post Local Similarity: 100.00; Prod. No. 24;
40 Matches: 7; Conservation: 0; Mismatches: 0; Indels: 0;
41 Gaps: 0;
42
43 QY: 58 LWILL: 64
44 ID: 7 LWILL: 14
45
46 RESULT:
47 US-08-474-040-21
48 Sequence 21, April 04/00 US/08474040
49 Patent No. 569372
50 GENERAL INFORMATION:
51 APPLICANT: GILLEN, Gary L.
52 APPLICANT: GILLEN, Gary L.
53 APPLICANT: SCHNEIDER, William P.
54 APPLICANT: LANE, Nicholas P.
55 APPLICANT: GILLEN, Kathleen E.
56 APPLICANT: SELLER, Harold E.
57 TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
58 NUMBER OF SEQUENCES: 114
59 CORRESPONDENCE ADDRESS:
60 ADDRESSEE: Townsend and Townsend and Crew
61 STREET: 879 Lydon Avenue
62 CITY: Palo Alto
63 STATE: California
64 COUNTRY: US
65 ZIP: 94301
66 COMPUTER READABLE FORM:
67 MEDIUM TYPE: Floppy disk
68 COMPUTER: IBM PC compatible
69 OPERATING SYSTEM: pc dos/MS-DOS
70

```

SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,487 200
 FILING DATE: 7-JUN-1995
 CLASSIFICATION: 4.24
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/644,278
 FILING DATE: 19-DEC-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/590,274
 FILING DATE: 28-SEP-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/710,252
 FILING DATE: 13-FEB-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/290,975
 FILING DATE: 28-DEC-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, William M
 REGISTRATION NUMBER: 30,224
 REFERENCE/POWER NUMBER: 11824-002610
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 926-2422
 TELEFAX: (415) 926-2422
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 126 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US 08/484,487 21

Query Match: 2.48, Score 7, DB 1, Length 126
 Best Local Similarity: 100.0%; Prod. No. 24;
 Matches: 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

QY 58 ILWILL 64
 ID 7 ILWILL 14

RESULT 10
 US-08-656-586 6
 Sequence 6, Application 05/0845576
 Patent No. 684597
 GENERAL INFORMATION:
 APPLICANT: Tso, J Yun
 APPLICANT: Goto, Michael S.
 TITLE OF INVENTION: Mutated No. 584597activation IgG2 domains and
 TITLE OF INVENTION: Anti CD4 Antibodies Incorporating IgG2
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-6444
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM pc compatible
 OPERATING SYSTEM: PC DOS/MS DOS
 SOFTWARE: Patent in Release #1.0, Version #1.00
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,487 6
 FILING DATE: 31-MAY-1996
 CLASSIFICATION: 5.80
 ATTORNEY/AGENT INFORMATION:
 NAME: Jacobsen, Joseph O.
 REGISTRATION NUMBER: 47,505
 REFERENCE/POWER NUMBER: 11824-0021005

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 976-0200
 TELEFAX: (415) 976-0400
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 126 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US 08 656 586 6

Query Match: 2.48, Score 7, DB 1, Length 126
 Best Local Similarity: 100.0%; Prod. No. 24;
 Matches: 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

QY 58 ILWILL 64
 ID 7 ILWILL 14

RESULT 11
 US-08-484-547-21
 Sequence 21, Application 05/0844547
 Patent No. 6180470
 GENERAL INFORMATION:
 APPLICANT: Queen, Gary L.
 APPLICANT: Goto, Man Sheng
 APPLICANT: Stambler, William P.
 APPLICANT: Landolfi, Nicholas P.
 APPLICANT: O'Leary, Kathleen L.
 APPLICANT: Selick, Harold E.
 TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
 NUMBER OF SEQUENCES: 114
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew
 STREET: 479 Lytton Avenue
 CITY: Palo Alto
 STATE: California
 COUNTRY: US
 ZIP: 94301
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM pc compatible
 OPERATING SYSTEM: PC DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,537
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/644,278
 FILING DATE: 19-DEC-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/590,274
 FILING DATE: 28-SEP-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/710,252
 FILING DATE: 13-FEB-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/290,975
 FILING DATE: 28-DEC-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, William M
 REGISTRATION NUMBER: 30,224
 REFERENCE/POWER NUMBER: 11824-002600
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 926-2400
 TELEFAX: (415) 926-2422
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 126 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

Query Match 2.38; Score 7; DB 4; Length 128;
 Best Local Similarity 100.08; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

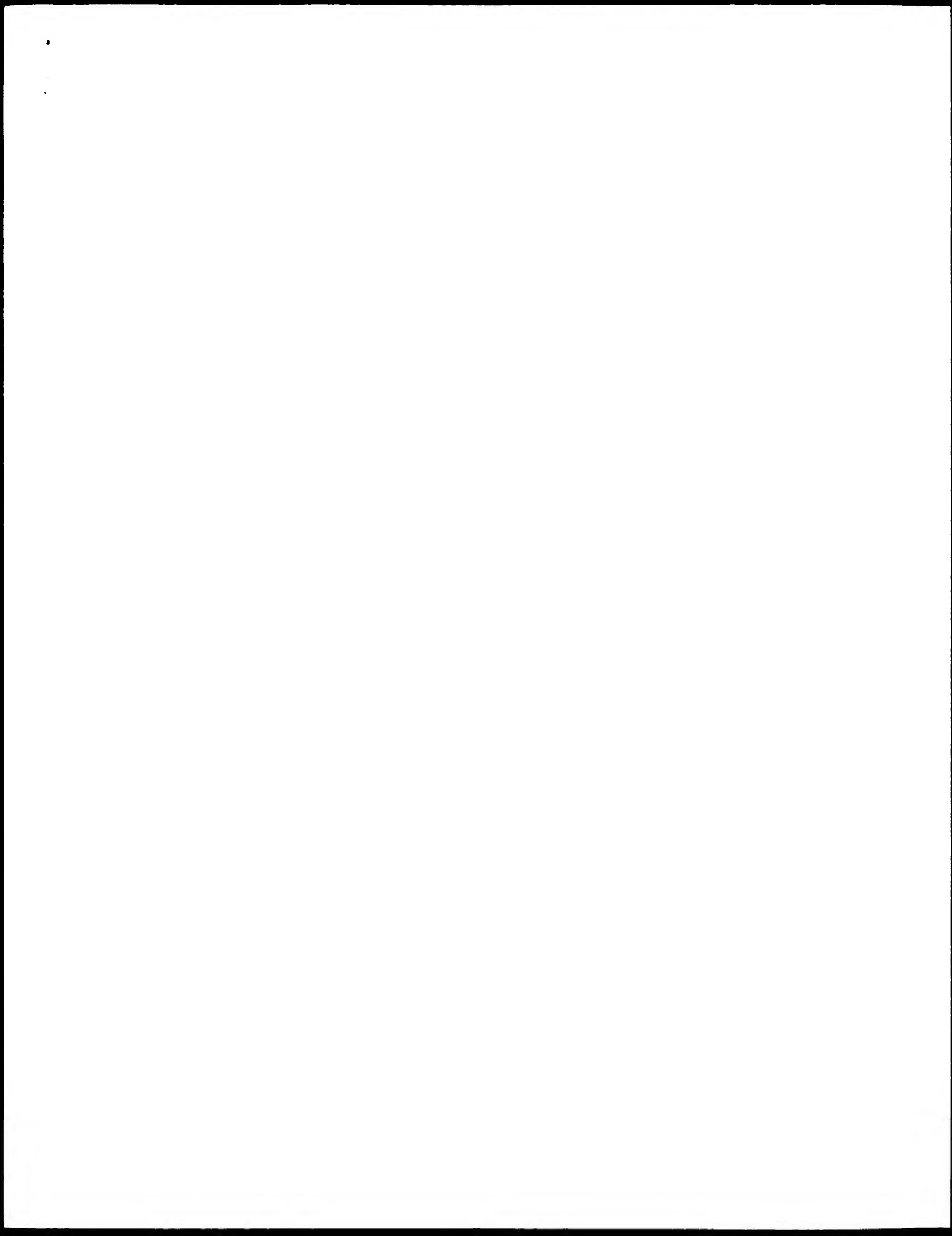
QY 58 L1WVILL.64
 |||||
 DB 7 L1WVILL.13

RESULT 15
 PCT-US95-15716-6
 : Sequence 6, Application PC/TUS9515716
 : GENERAL INFORMATION:
 : APPLICANT: Berdoz, Jose
 : APPLICANT: Kichenbuhl, Jean Pierre
 : TITLE OF INVENTION: PCR AMPLIFICATION OF PEAPANDED GENOMIC
 : TITLE OF INVENTION: VARIABLE REGIONS OF IMMUNOGLOBULIN GENES
 : NUMBER OF SEQUENCES: 108
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Fish & Richardson
 : STREET: 225 Franklin Street, Suite 3100
 : CITY: Boston
 : STATE: MA
 : COUNTRY: USA
 : ZIP: 02110-2804
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.30B
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US95/15716
 : FILING DATE:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/348,548
 : FILING DATE: 01-DEC-1994
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Clark, Paul T.
 : REGISTRATION NUMBER: 30,162
 : REFERENCE/BOOKET NUMBER: 06132/000001
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (617) 542-5070
 : TELEFAX: (617) 542-5070
 : TELEX: 200154
 : INFORMATION FOR SEQ. 1: NO: 6;
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 128 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: not relevant
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : PCT US95-15716-6

Query Match 2.38; Score 7; DB 5; Length 128;
 Best Local Similarity 100.08; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 L1WVILL.64
 |||||
 DB 7 L1WVILL.13

Search completed: August 12, 2002, 14:00:02
 Job time: 249 sec



1766416 (1766416:2001-01-17&4-1)

Seq primer: puc 18 forward
High quality sequence starts 4
High quality sequence stops 236
Location/Qualifiers

FEATURES

SOURCE

1..236

/organism "Homo sapiens"

/db_xref "taxon:9606"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

FEATURES

SOURCE

1..236

/organism "Homo sapiens"

/db_xref "taxon:9606"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

/clone_11b "11b008"

Tue Aug 13 08:15:38 2002

us-09-825-882-7.oli.rst

Page 8

Search completed: August 12, 2002, 14:11:54
1 to 4 lines, 196 lines